Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

GKI-Net: Gaussian Splatting and KAN-based Implicit Neural Representation for Physically Grounded Dehazing

Xiaoji Wei
Jiaxing Vocational and Technical College
Zhongbin Luo
luozhongbin@stu.cqu.edu.cn
Chongqing University
Bilu Luo
Jiaxing Vocational and Technical College
Yuhan Chen
Chongqing University

Research Article

Keywords: Image Dehazing, Gaussian Splatting, Kolmogorov-Arnold Networks

Posted Date: June 8th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-8958794/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Version of Record: A version of this preprint was published at Signal, Image and Video Processing on October 20th, 2026. See the published version at https://doi.org/10.1007/s11760-026-05771-y.

GKI-Net: Gaussian Splatting and KAN-based Implicit Neural Representation for Physically Grounded Dehazing

Xiaoji Wei[1], Zhongbin Luo[2,3*], Bilu Luo[1], Yuhan Chen[4]

[1]Jiaxing Vocational and Technical College, Jiaxing, 314036, China.
[2*]College of Computer Science, Chongqing University, Chongqing, 400044, China.
[3]China Merchants Chongqing Communications Research and Design Institute Co., Ltd., Chongqing, 400067, China.
[4]College of Mechanical and Vehicle Engineering, Chongqing University, Chongqing, 400044, China.

*Corresponding author(s). E-mail(s): luozhongbin@stu.cqu.edu.cn;
Contributing authors: weixiaoji@jxvtc.edu.cn; luobilu@jxvtc.edu.cn;
20240701028@stu.cqu.edu.cn;

Abstract

Haze causes severe image degradation in visual sensing and obscures key physical cues. We propose GKI-Net, a physics-embedded dehazing framework that integrates Gaussian splatting with a convolutional Kolmogorov–Arnold network. Instead of relying on black-box mappings, we design a hybrid physics-informed decoder. Using multi-scale implicit neural representations, the decoder explicitly decouples the transmission map and atmospheric light in the atmospheric scattering model, thereby preserving physical consistency in the restored images. During feature encoding, we design a Gaussian feature splatting module that uses deformable Gaussian kernels to adaptively aggregate long-range spatial context. We further introduce a convolutional KAN layer (ConvKAN) to perform efficient high-order nonlinear feature transformations with a small number of parameters. Experiments show that GKI-Net outperforms state-of-the-art methods on quantitative metrics and visual quality, improving the robustness of sensing systems in complex weather conditions.

Keywords: Image Dehazing, Gaussian Splatting, Kolmogorov-Arnold Networks

1 Introduction

The rapid growth of the Internet of Things and intelligent surveillance has made visual sensors a key perception component in autonomous driving, drone inspection, and security monitoring. In real-world deployments, these sensors are often affected by adverse weather, especially haze. Suspended particles in the atmosphere scatter and absorb light, leading to reduced contrast, color shifts, and loss of fine details in captured images [1, 2]. This loss in image quality reduces visual clarity and can substantially degrade downstream vision tasks such as object detection and semantic segmentation. As a result, it poses serious challenges to the safety and reliability of vision-based intelligent systems [3].

Efficient single-image dehazing methods that are physically interpretable and robust in complex scenes are therefore important for improving the resilience of visual sensing systems in real-world conditions. Over the past decades, image dehazing has shifted from hand-crafted approaches built on physical priors to data-driven methods based on deep learning. Early approaches mainly followed the atmospheric scattering model (ASM), which represents a hazy image as a linear combination of the transmission map and atmospheric light. By exploiting the statistical characteristics of clear images, several priors have been introduced, such as the Dark Channel Prior (DCP) [4] and the Color Attenuation Prior (CAP) [5]. Although these prior-based methods perform effectively in specific scenarios, their underlying assumptions often become invalid in sky regions or on bright surfaces, leading to color distortions and halo artifacts in the restored results.

With the advent of Convolutional Neural Networks (CNNs), deep learning has become the dominant paradigm in image dehazing. Pioneering works like DehazeNet [6] and AOD-Net [7] employed neural networks to estimate physical parameters, enabling end-to-end restoration. Subsequent studies integrated multi-scale fusion [8], attention mechanisms [9, 10], and Generative Adversarial Networks (GANs) [11, 12] to enhance feature representation capabilities. For instance, FFA-Net [9] significantly enhanced performance by incorporating a feature fusion attention module, whereas PSD [13] leveraged physical priors to bridge the domain gap between synthetic and real-world data. While recent diffusion-based models [14, 15] excel in generating high-fidelity textures, their prohibitive computational costs and high inference latency hinder their deployment in real-time visual sensing scenarios.

Crucially, many existing deep dehazing networks establish a direct "black-box" mapping from hazy to haze-free images, thereby neglecting the underlying physical formation mechanism of atmospheric scattering. This absence of physical grounding compromises model interpretability and frequently generates artifacts that violate physical laws, manifesting as inconsistencies in depth and illumination. To address these limitations, it is imperative to explore a feature representation that efficiently captures long-range spatial contexts while preserving physical consistency.

Recently, 3D Gaussian Splatting (3DGS) [16] has emerged as a prominent explicit neural representation, demonstrating exceptional speed and reconstruction quality in 3D scenes. In contrast to implicit neural representations like NeRF, 3DGS characterizes scene attributes using anisotropic Gaussian ellipsoids. Recent studies indicate

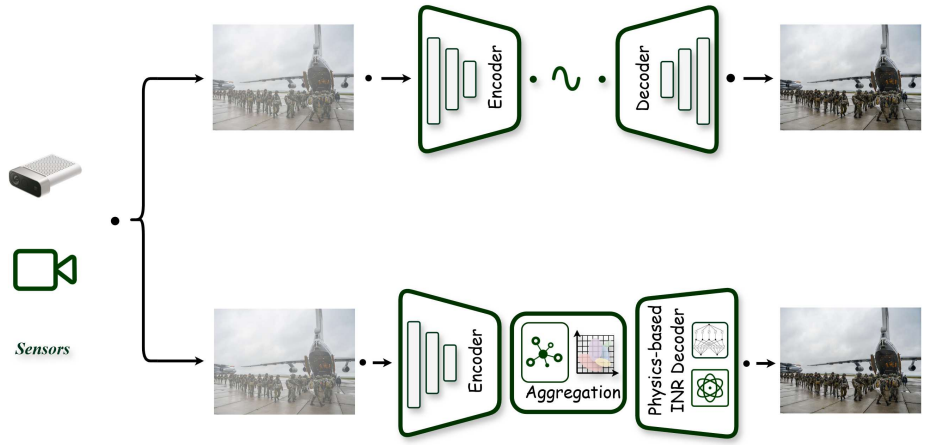


Fig. 1 (A) illustrates conventional dehazing frameworks. Most of these methods adopt U-Net or coarse-to-fine architectures as the primary backbone. They take hazy images as input and execute a direct end-to-end mapping driven purely by data within a black-box neural network. (B) depicts the proposed method. We construct deformable Gaussian kernels from the encoded features to adaptively aggregate long-range spatial context instead of relying on traditional local convolutions. Furthermore, we integrate Convolutional KAN layers into a hybrid physics decoder to enforce explicit physical parameter decoupling.

that 2D Gaussian Splatting (2DGS) functions as a highly efficient image representation [17, 18]. Notably, GaussianImage [18] demonstrates that 2DGS kernels can approximate high-frequency details with minimal parameters. While existing works have applied 3DGS to multi-view dehazing [19–22], the potential of 2DGS in single-image dehazing remains largely untapped. Specifically, leveraging the deformable properties of 2DGS kernels to model spatially non-uniform haze distributions and embedding this representation into a physics-based restoration framework constitutes an underexplored avenue.

Furthermore, traditional Convolutional Neural Networks (CNNs) and Multilayer Perceptrons (MLPs) exhibit limited efficiency in capturing high-frequency nonlinear features. Conversely, the recently proposed Kolmogorov-Arnold Networks (KANs) present a spline-based alternative, achieving superior nonlinear mapping capabilities with significantly fewer parameters.

To address these challenges, we propose GKI-Net, a physics-embedded dehazing network integrating Gaussian Splatting with Convolutional Kolmogorov-Arnold Networks. As shown in Fig. 1, diverging from conventional approaches that expand receptive fields through the extensive stacking of convolutional layers, GKI-Net innovatively incorporates a Gaussian Feature Splatting Module (GFSM). By leveraging learnable offsets and scale parameters to adaptively modulate the shape and position of Gaussian kernels, this module facilitates long-range, irregular context aggregation within the feature space. This mechanism inherently aligns with the physical characteristics of spatially non-uniform atmospheric haze distribution.

To simultaneously enhance nonlinear feature representation while maintaining a lightweight architecture, we integrate efficient Convolutional KAN (ConvKAN) layers into the decoder. Crucially, GKI-Net avoids direct black-box mappings by employing a Hybrid Physical Implicit Decoder (HPD). This decoder explicitly decouples and estimates the transmission map t and atmospheric light A, enabling the reconstruction of the haze-free image J via the atmospheric scattering model. This design ensures physical consistency in the restoration and provides transparent, visualizable intermediate variables. The main contributions of this work are summarized as follows:

1. We propose GKI-Net, a physics-embedded dehazing framework tailored for visual sensors. It is the first to combine 2D Gaussian splatting with a convolutional Kolmogorov–Arnold network, alleviating the limits of conventional CNNs in long-range dependency modeling and parameter efficiency. The restoration process follows the atmospheric scattering model to maintain physical fidelity.

2. We design a Gaussian Feature Splatting Module (GFSM) with deformable Gaussian kernels that adaptively aggregate spatial information based on local feature richness. Visualizations show that GFSM concentrates on dense haze and edge regions, improving the robustness of feature extraction.

3. By combining ConvKAN for efficient nonlinear mapping with the physical constraints in HPD, GKI-Net explicitly estimates atmospheric parameters while producing high-quality restored images.

2 Related Work

2.1 Evolution of Single Image Dehazing Algorithms

Image dehazing constitutes a pivotal low-level vision task aimed at recovering haze-free scene radiance from degraded observations. Based on the dependence on atmospheric scattering models and optimization strategies, existing methodologies have progressed through three distinct phases: handcrafted priors, supervised learning, and unsupervised or generative learning paradigms.

Prior-Based and Supervised Learning Methods: Early dehazing approaches predominantly relied on statistical observations of natural images, employing handcrafted priors to estimate the unknown variables in the atmospheric scattering model, specifically the transmission map t and atmospheric light A. A seminal work in this domain is the Dark Channel Prior (DCP) proposed by He et al. [4], which posits that at least one color channel within non-sky local patches exhibits an intensity approaching zero. Zhu et al. introduced the color attenuation prior (CAP) [5], which estimates haze density by modeling a linear relationship between scene depth and image brightness and saturation. However, these priors often break down in cases that deviate from the assumed statistics, such as white objects or overexposed regions, leading to color distortions and visible artifacts.

The emergence of Convolutional Neural Networks has established data-driven approaches as the dominant paradigm. Pioneering works such as DehazeNet [6] and MSCNN [8] utilized CNNs to estimate transmission maps directly. Subsequently, AOD-Net [7] reformulated the atmospheric scattering model into an end-to-end

joint optimization framework, effectively circumventing the cumulative errors inherent in multi-stage estimation strategies. To leverage multi-scale features effectively, GridDehazeNet [10] proposes a grid-based architecture that facilitates information propagation. Capitalizing on attention mechanisms, FFA-Net [9] employs Feature Fusion Attention modules to substantially enhance the restoration of high-frequency details. Furthermore, Enhanced pix2pix [23] and PSD [13] investigate the boundaries of dehazing performance through adversarial learning and synthetic-to-real transfer paradigms, respectively.

Unsupervised, Generative and Emerging Paradigms: To circumvent the scarcity of paired real-world dehazing datasets, subsequent methodologies extensively employ unsupervised learning and Generative Adversarial Networks (GANs). CycleGAN [24] and its variant Cycle-Dehaze [12] leverage cycle-consistency constraints to achieve domain-adaptive dehazing on unpaired data. Furthermore, Yang et al. [25] and Li et al. [26] introduce feature decoupling concepts to isolate haze representations from underlying image content. RefineDNet [27] and Zero-shot dehazing [28] investigate self-supervised or weakly supervised strategies in the absence of reference images.

More recently, approaches such as Contrastive Learning [29], UCL-dehaze [30], and Self-augmented dehazing [31] leverage contrastive learning paradigms to enhance feature discriminability. These methods improve restoration quality by minimizing the feature-space distance between dehazed outputs and clear images. Recent academic focus has increasingly gravitated towards Generative Diffusion Models. Lan et al. [14] leverage diffusion priors to tackle unpaired dehazing tasks, while Liu et al. [15] introduce a frequency-domain diffusion model to expedite convergence. Concurrently, the Schrodinger Bridge framework [32] establishes an alternative probabilistic pathway for transitioning from noise to haze-free images. Furthermore, HardGAN [11] enhances GAN-based generation quality by integrating haze-aware representation distillation.

Motivation for GKI-Net: Despite significant advancements in existing methodologies, two critical limitations persist. First, predominant deep learning frameworks, including CNNs and Transformers, often lack explicit physical interpretability when approximating the nonlinear mappings inherent to atmospheric scattering models. Second, standard convolutional operations struggle to model globally distributed degradations like haze due to their local receptive fields, thereby failing to capture essential long-range spatial dependencies. To address these limitations, GKI-Net departs from conventional black-box mappings by adopting a physics-embedded strategy. We leverage the efficient nonlinear approximation capabilities of ConvKAN to explicitly decouple the transmission map and atmospheric light. Furthermore, by incorporating implicit neural representations, our method ensures physical self-consistency in the restored results, yielding robust dehazing performance in complex, dynamic environments.

2.2 Gaussian Splatting: From 3D Rendering to 2D Representation

The Revolution of 3D Gaussian Splatting: Kerbl et al. [16] introduced 3D Gaussian Splatting (3DGS) to represent scenes as a set of anisotropic 3D Gaussians. By integrating a tile-based rasterizer, this approach achieves real-time rendering at megapixel

resolutions. In contrast to implicit Neural Radiance Fields (NeRFs), 3DGS leverages an explicit point-based structure, facilitating rapid training convergence. This line of work has quickly expanded to more complex settings. Huang et al. [33] and Hanson et al. [34] improved the geometric accuracy and rendering speed of Gaussian representations.

For autonomous driving and dynamic scenes, Street Gaussians [35], DrivingGaussian [36], GaussianPro [37], and CityGaussian [38] tackle dynamic object modeling in large-scale urban environments. Momentum-GS [39] improves reconstruction quality through momentum self-distillation. PMGS [40] and PEGS [41] further incorporate physical constraints to reconstruct parabolic motion over large spatiotemporal ranges. For generative tasks, GaussianDreamer [42] and ReconDreamer [43] combine diffusion models with 3D Gaussians for text-to-3D generation and scene completion.

Gaussian Splatting in Hazy Scenes: Given the advantages of 3D Gaussian splatting in structural representation, it has been explored for reconstruction in hazy scenes. DehazeGS [19, 20] and DehazeSplat [21] support joint dehazing and 3D reconstruction by introducing haze-related parameters into 3D Gaussian attributes or by using multi-view geometric constraints. SmokeSeer [22] addresses scene restoration and reconstruction within smoky environments. Nevertheless, these approaches primarily target multi-view 3D reconstruction and require geometric consistency across multiple viewpoints. Their substantial computational overhead hinders direct application to real-time single-image dehazing tasks.

2D Gaussian Image Representation: More recently, the core principles of Gaussian splatting have transitioned towards 2D image representations. GaussianImage [18] serves as a pioneering framework within this paradigm and has inspired a wide range of derivative methodologies. Following this direction, Zhu et al. [44] utilize 2D Gaussian primitives to achieve efficient fitting of large-scale image data, demonstrating the scalability of this representation. Instant GaussianImage [45] introduces an adaptive image representation scheme. Meanwhile, Jiang et al. [46] leverage sparse Gaussian representations to facilitate efficient dataset distillation.

Furthermore, this paradigm expands into various low-level vision tasks, including image compression [47], super-resolution [48], and low-light enhancement [17, 49]. These methodologies generally employ 2D Gaussians as rendering primitives, approximating pixel values by optimizing the positions, colors, and covariances of the Gaussians.

Motivation for GKI-Net: Unlike existing Gaussian image methods that limit Gaussians to pixel-level rendering units, we introduce the Gaussian Feature Splatting Module (GFSM). We observe that haze distributions exhibit continuity across large scales and stochasticity within fine texture details. While standard convolutional kernels maintain fixed geometries, Gaussian kernels are inherently deformable, utilizing the covariance matrix to adaptively control orientation and scale. Rather than rendering colors directly, GKI-Net employs learnable Gaussian kernels to perform long-range context aggregation within the feature space. By estimating the Gaussian offset, variance, and amplitude for each feature point, our framework adaptively perceives non-uniform haze distributions to facilitate content-aware feature extraction. This

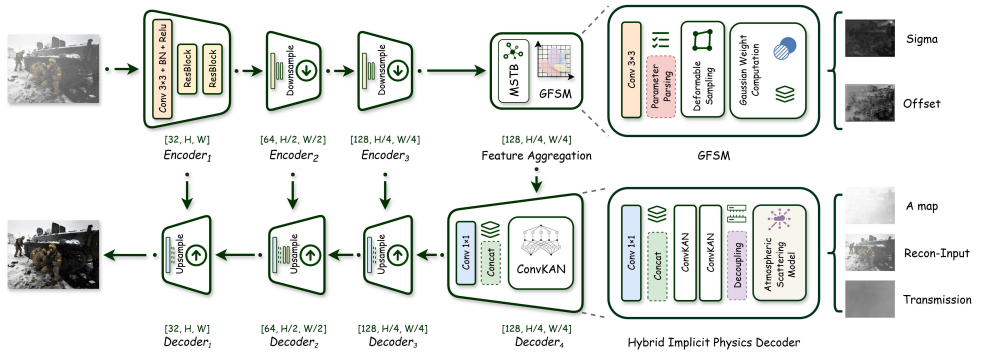


Fig. 2 The overall architecture of GKI-Net. The input hazy image is processed by a multi-scale encoder to extract hierarchical features, followed by the Gaussian Feature Splatting Module (GFSM) to adaptively aggregate long-range spatial contexts. Subsequently, the physics-guided hybrid physical decoder explicitly decouples the transmission map t, atmospheric light A, and scene radiance J across multiple scales. It utilizes the Atmospheric Scattering Model $I(x) = J(x)t(x) + A(1 - t(x))$ to enforce physical consistency constraints, ensuring the interpretability and robustness of the restoration results.

architecture preserves the spatial modeling flexibility of Gaussian splatting while overcoming the receptive field constraints of conventional CNNs when processing global degradations.

3 Proposed Method

The primary challenge in image dehazing stems from its nature as an ill-posed inverse problem. To regularize the solution space using physical constraints within a deep learning framework, we first revisit the classical Atmospheric Scattering Model (ASM). During the degradation process, the sensor-captured irradiance $I(x)$ is defined as a linear combination of the haze-free radiance $J(x)$ and the global atmospheric light A, where the medium transmission $t(x)$ determines the respective weights. The ASM is formulated as

$$I(x) = J(x)t(x) + A(1 - t(x)) \tag{1}$$

Where $x = (u, v)$ represents the pixel-wise spatial coordinates.

The transmission map $t(x) \in (0, 1]$ characterizes the capacity of light to penetrate the medium and exhibits an exponential decay relative to the scene depth $d(x)$ and the atmospheric scattering coefficient β:

$$t(x) = e^{-\beta d(x)} \tag{2}$$

Standard dehazing frameworks primarily utilize deep neural networks to establish end-to-end black-box mappings from $I(x)$ to $J(x)$. However, these approaches often bypass the physical formation mechanisms of haze, leading to suboptimal generalization under non-uniform haze or extreme illumination. To overcome these limitations, we introduce GKI-Net. As illustrated in Fig. 2, GKI-Net employs a multi-scale autoencoder architecture designed to facilitate high-quality image restoration through the

explicit decoupling of physical parameters. Our approach incorporates three core innovations:

1. *ConvKAN-based Feature Transformation:* We utilize the approximation power of B-splines to substitute conventional MLPs. This design facilitates high-order nonlinear feature mapping with enhanced parameter efficiency.

2. *Physics-Embedded Gaussian Feature Splatting Module:* This module employs deformable Gaussian kernels for feature resampling and aggregation. It effectively emulates the non-local propagation characteristics of light within scattering media.

3. *Hybrid implicit physical representation decoder:* By integrating spatial coordinates, this decoder explicitly maps latent features into physical parameters J, t, and A while enforcing physical consistency constraints.

Network Architecture: Given an input hazy image $I \in \mathbb{R}^{3\times H\times W}$, an encoder built from residual blocks first extracts multi-scale features F_s. The features then pass through the multi-scale topology hybrid module (MSTB) at the bottleneck and the GFSM for global context aggregation. Finally, the decoder predicts the physical parameters at three scales, forming a coarse-to-fine restoration process.

3.1 Convolutional Kolmogorov-Arnold Networks

Conventional Convolutional Neural Networks (CNNs) operate via linear convolution operations followed by fixed nonlinear activation functions. According to the Kolmogorov-Arnold representation theorem, arbitrary multivariate continuous functions can be represented as a finite composition of univariate nonlinear functions. Drawing inspiration from this, we introduce the Convolutional KAN (ConvKAN) layer for feature extraction. This design positions learnable activation functions on the connection weights rather than at the neuron nodes.

Base Transformation and Spline-Basis Expansion: For an input feature tensor $X \in \mathbb{R}^{C_{\text{in}}\times H\times W}$, we first apply Instance Normalization followed by a hyperbolic tangent function (Tanh). This operation stabilizes the feature distribution and maps the input to the effective B-spline domain, defined as $(-1, 1)$ in our setting:

$$\tilde{X} = \text{Tanh}(\text{InstanceNorm}(X)) \tag{3}$$

The core computation of ConvKAN comprises two parallel pathways: a basis function branch and a spline branch. The basis function branch resembles a standard convolutional layer, designed to capture linear mapping relationships within the input features while utilizing the SiLU activation function to introduce basic nonlinearity. Given the linear transformation weights W_{base}, we formulate the output Y_{base} as

$$Y_{\text{base}} = W_{\text{base}} * \text{SiLU}(X) \tag{4}$$

where $*$ represents the convolution operation. To maintain variance stability, we scale W_{base} using Kaiming initialization.

In the spline branch, input features are projected onto a set of B-spline basis functions. Assuming a spline order k and G grid intervals, we establish uniform grid points $\{g_i\}$ over the domain $[x_{\min}, x_{\max}]$. For an input scalar x, the i-th B-spline basis

function $B_{i,k}(x)$ follows the Cox–de Boor recursion formula:

$$\begin{cases} B_{i,0}(x) = \begin{cases} 1 & \text{if } g_i \leq x < g_{i+1} \\ 0 & \text{otherwise} \end{cases} \\ B_{i,k}(x) = \dfrac{x - g_i}{g_{i+k} - g_i} B_{i,k-1}(x) + \dfrac{g_{i+k+1} - x}{g_{i+k+1} - g_{i+1}} B_{i+1,k-1}(x) \end{cases} \tag{5}$$

In the vectorized implementation, we employ tensor expansion to map the input X into a basis function space with dimensions $C_{\text{in}} \times (G + k)$. Let $B(x)$ represent the vector comprising all basis functions. The output of the spline branch Y_{spline} is formulated as a linear combination of basis responses:

$$Y_{\text{spline}} = \sum_{j=1}^{C_{\text{in}}} \sum_{m=1}^{G+k} w_{c,j,m} \cdot B_{m,k}(X_j) \tag{6}$$

where $w_{c,j,m}$ denotes the learnable spline control coefficients. Computationally, we optimize this step as an efficient matrix multiplication by flattening the features followed by a linear transformation.

Aggregation and Output: We compute the final ConvKAN output by summing the responses from the basis and spline branches:

$$Y_{\text{out}} = Y_{\text{spline}} + Y_{\text{base}} \tag{7}$$

This architecture preserves the advantages of conventional convolutions for high-frequency texture extraction while leveraging the B-splines' robust capability to fit complex low-frequency nonlinear mappings. Relative to standard MLPs, ConvKAN achieves superior fitting accuracy with reduced parameter overhead. This efficiency is critical for recovering missing textural details in dehazing tasks.

3.2 Physically Grounded Gaussian Feature Splatting

Haze distribution typically exhibits spatial non-uniformity, while atmospheric scattering is characterized by long-range dependencies. Standard convolution kernels, restricted by local receptive fields, often fail to effectively model these global degradation patterns. Drawing inspiration from the expressive capability of 3D Gaussian Splatting in novel view synthesis, we propose the Gaussian Feature Splatting Module (GFSM). This module conceptualizes image features as a "feature cloud" and employs deformable Gaussian kernels to execute adaptive aggregation within the feature space.

Gaussian Parameter Regression: Given the input feature map $F_{\text{in}} \in \mathbb{R}^{C \times H \times W}$, we employ a lightweight convolutional network to predict Gaussian distribution parameters at each pixel coordinate $p = (u, v)$. The network generates a tensor with $5 \times K^2$ channels, corresponding to the sampling points within the Gaussian kernel. Specifically, for the k-th sampling point, we regress the spatial offset Δp_k, covariance σ_k, and amplitude α_k. To guarantee numerical stability and physical validity, we apply

specific activation functions to the raw outputs:

$$\begin{cases} \Delta p_k = \text{Tanh}(O_{\text{pos}}) \cdot \dfrac{K}{2} \\ \sigma_k = \text{Softplus}(O_{\text{scale}}) + \epsilon \\ \alpha_k = \text{Sigmoid}(O_{\text{amp}}) \end{cases} \tag{8}$$

where $\epsilon = 0.1$ represents a small constant introduced to prevent division by zero.

Deformable Sampling and Gaussian Weighting: Upon obtaining the offset Δp_k, we determine the absolute sampling coordinate p'_k for the k-th sampling point. Given the current pixel's center p_0 and the regular grid offset $\delta_{\text{grid},k}$, we compute the location as:

$$p'_k = p_0 + \delta_{\text{grid},k} + \Delta p_k \tag{9}$$

We employ bilinear interpolation to retrieve the feature value $F(p'_k)$ at the non-integer coordinate p'_k. This strategy enables the network to transcend fixed-grid constraints and adaptively capture long-range contextual information. Subsequently, we derive the Gaussian weight for each sampling point based on the predicted variances $\sigma_{u,k}, \sigma_{v,k}$ and amplitude α_k. Defining the offset components as $\Delta p_k = (\delta u_k, \delta v_k)$, we formulate the unnormalized Gaussian response G_k as:

$$\begin{cases} D_k^2 = \|\Delta p_k\|^2 = \delta u_k^2 + \delta v_k^2 \\ E_k = -\left(\dfrac{\delta u_k^2}{2\sigma_{u,k}^2} + \dfrac{\delta v_k^2}{2\sigma_{v,k}^2} \right) \\ G_k = \alpha_k \cdot \exp(E_k) \end{cases} \tag{10}$$

To guarantee energy conservation and maintain numerical stability, we normalize the weights across the kernel dimension K^2:

$$W_k = \frac{G_k}{\sum_{j=1}^{K^2} G_j + \epsilon} \tag{11}$$

Feature Aggregation: We compute the final GFSM output F_{out} by aggregating the weighted sampled features combined with a residual connection:

$$\begin{cases} F_{\text{agg}}(p) = \displaystyle\sum_{k=1}^{K^2} W_k(p) \cdot F(p'_k) \\ F_{\text{out}} = \text{Proj}(F_{\text{agg}}) + F_{\text{in}} \end{cases} \tag{12}$$

Conceptually, this operation functions as a physics-guided attention mechanism. The spatial offsets Δp enable the network to identify relevant context, such as texture cues from haze-free regions. Meanwhile, the variance σ emulates the point spread function of the scattering medium, and the amplitude α dynamically modulates the fusion intensity.

3.3 Hybrid Implicit Physics Decoder

To estimate the physical quantities governing haze degradation, we introduce the Hybrid Implicit Physical Representation Decoder, diverging from standard convolutional decoders. Leveraging implicit neural representations, this module explicitly utilizes spatial coordinates as inputs to model the smooth variations of transmission maps and atmospheric light within a continuous domain.

Fusion of Coordinates and Features: Given the decoder input feature map $F_{\text{dec}} \in \mathbb{R}^{C\times H\times W}$, we construct a normalized spatial coordinate grid $C_{\text{coord}} \in \mathbb{R}^{2\times H\times W}$ with values spanning $[-1, 1]$. We concatenate this coordinate grid with the image features along the channel dimension to generate the hybrid representation Z:

$$Z = \text{Concat}(F_{\text{dec}}, C_{\text{coord}}) \tag{13}$$

Integrating absolute spatial coordinates enables the network to encode positional dependence. This mechanism is critical for accurately inferring atmospheric light A and transmission map t, which typically exhibit smooth variations correlated with depth.

Physical Parameter Decoupling: We feed the hybrid feature Z into a sequence of ConvKAN layers or residual blocks for deep feature transformation, generating the latent variable L. Subsequently, three independent projection heads decouple the physical components: the transmission map t, the atmospheric light map A, and the clear image residual J_{res}. We formulate this decomposition as:

$$\begin{cases} t = \text{Sigmoid}(\text{Head}_t(L)) \\ A = \text{Sigmoid}(\text{Head}_A(L)) \\ J_{\text{res}} = \text{Head}_J(L) \end{cases} \tag{14}$$

To accelerate convergence and leverage input information, we add the downsampled hazy image I_{down} to J_{res}. We derive the final restored image J by clamping the result to the valid intensity range:

$$J = \text{Clamp}(J_{\text{res}} + I_{\text{down}}, 0, 1) \tag{15}$$

Physically Consistent Reconstruction: To enforce adherence to the atmospheric scattering model, we introduce a self-supervised physical reconstruction constraint during training. Utilizing the predicted components J, t, and A, we reconstruct the hazy image I_{recon} based on (1):

$$I_{\text{recon}} = J \cdot t + A \cdot (1 - t) \tag{16}$$

Minimizing the discrepancy between I_{recon} and the original input I guarantees that the decomposed physical parameters mathematically align with the governing laws of haze formation.

3.4 Loss Functions

To comprehensively supervise GKI-Net training, we design a composite loss function L_{total} comprising five terms: Charbonnier reconstruction, edge fidelity, physical consistency, total variation, and dark channel prior constraints.

We primarily employ the Charbonnier loss L_{char} for pixel-wise supervision. Relative to standard L_2 metrics, this variant demonstrates superior robustness against outliers and preserves high-frequency details more effectively. Given the predicted clear image J and ground truth J_{gt}, we calculate:

$$L_{\text{char}}(J, J_{\text{gt}}) = \sqrt{\|J - J_{\text{gt}}\|^2 + \epsilon^2} \tag{17}$$

where the constant ϵ is set to 10^{-3}. We compute this reconstruction loss across three scales:

$$L_{\text{rec}} = \sum_{s \in \{1, 1/2, 1/4\}} L_{\text{char}}(J_s, J_{\text{gt},s}) \tag{18}$$

To enhance high-frequency texture recovery, we introduce the edge fidelity loss L_{edge}. We utilize the Laplacian operator $\Delta(\cdot)$ to extract edge maps and minimize the L_1 difference between the predicted and ground-truth edges:

$$L_{\text{edge}} = \|\Delta(J) - \Delta(J_{\text{gt}})\|_1 \tag{19}$$

To ensure the reconstructed hazy image aligns with the input I, we enforce the physical consistency loss L_{phy}:

$$L_{\text{phy}} = \sum_s \|I_{\text{recon},s} - I_s\|_1 \tag{20}$$

Given that transmission maps are typically locally piecewise smooth, we impose total variation regularization on t to suppress noise artifacts:

$$L_{\text{tv}} = \sum_x (\nabla_u t(x)^2 + \nabla_v t(x)^2) \tag{21}$$

To enforce conformity with natural image statistics, particularly that the dark channel of haze-free regions approaches zero, we incorporate the dark channel loss L_{dc}. We define the dark channel J^{dark} and the corresponding loss as:

$$\begin{cases} J^{\text{dark}}(x) = \min\limits_{c \in \{r,g,b\}} \left(\min\limits_{y \in \Omega(x)} J^c(y) \right) \\ L_{\text{dc}} = \dfrac{1}{HW} \sum\limits_x \|J^{\text{dark}}(x)\|_1 \end{cases} \tag{22}$$

To prevent abrupt spatial fluctuations in the estimated atmospheric light A, we impose a smoothness constraint $L_{\text{smooth_A}}$ on its gradients:

$$L_{\text{smooth_A}} = \|\nabla A\|_1 \tag{23}$$

Consequently, we formulate the final optimization objective as a weighted summation of these components:

$$\begin{aligned} L_{\text{total}} = L_{\text{rec}} + \lambda_1 L_{\text{edge}} + \lambda_2 L_{\text{phy}} \\ + \lambda_3 L_{\text{tv}} + \lambda_4 L_{\text{dc}} + \lambda_5 L_{\text{smooth_A}} \end{aligned} \tag{24}$$

In our experiments, we empirically set the weighting hyperparameters to $\lambda_1 = 0.1$, $\lambda_2 = 0.1$, $\lambda_3 = 0.01$, $\lambda_4 = 0.001$, and $\lambda_5 = 0.01$. This joint optimization strategy enables GKI-Net to effectively eliminate haze while maximizing the preservation of structural details and color fidelity.

4 Experiments

4.1 Experimental Setup

Dataset: To comprehensively evaluate GKI-Net across dimensions of objective physical consistency and subjective visual perception, we select four representative dehazing datasets and categorize them into two groups. The first group comprises SOTS [50] and NID [51]. Containing paired hazy inputs and clear reference images, these datasets facilitate full-reference evaluation to quantify the accuracy of pixel-level reconstruction and structural recovery. The second group comprises RTTS [52] and Haze2020 [53]. Characterized by complex real-world hazy scenes, these datasets lack corresponding haze-free ground truths. We employ them for no-reference evaluation to verify the generalization capability of our method under real-world degradations and the perceptual naturalness of the restored images.

Implementation Details: We employ ConvKAN as the core unit for both the implicit physical representation and the decoder. We configure the spline interpolation order to 3 and the grid size to 3. This setup achieves efficient high-order nonlinear fitting while maintaining a low parameter count. Furthermore, the Gaussian Feature Splatting Module utilizes a 3×3 adaptive kernel to capture long-range spatial context.

We implement GKI-Net using the PyTorch framework and conduct end-to-end training and inference on a single NVIDIA RTX 3090 GPU. To guarantee experimental reproducibility, we provide detailed parameter configurations and training strategies. We train the model for 200 epochs with a batch size of 4. We utilize the AdamW optimizer initialized with a learning rate of 2×10^{-4} and a weight decay of 1×10^{-4}. We employ a cosine annealing strategy to smoothly decay the learning rate from the initial value to a minimum of 1×10^{-6}.

Evaluation Metrics: We employ two evaluation protocols tailored to specific data types. For the SOTS and NID datasets containing paired reference images, we select two full-reference metrics: PSNR and SSIM. PSNR quantifies the pixel-level fidelity

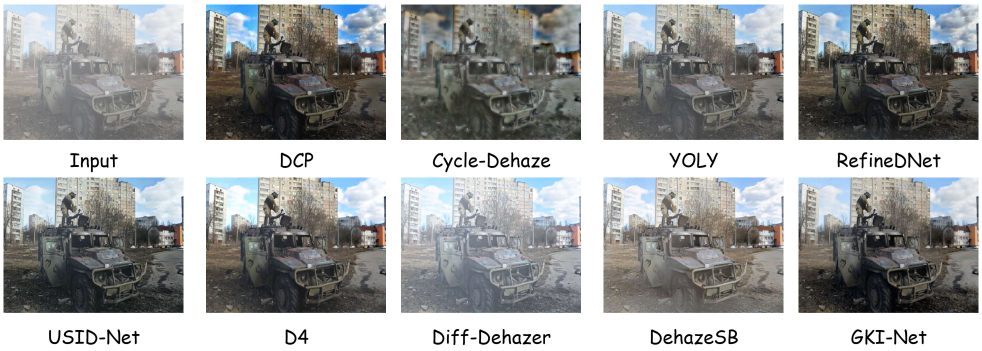


Fig. 3 Visual comparison of GKI-Net with state-of-the-art methods on the NID dataset.

of the dehazed outputs. SSIM evaluates structural similarity across luminance, contrast, and structure dimensions. For the RTTS and Haze2020 datasets representing real-world scenes without reference images, we utilize four advanced no-reference metrics: FID, NIQE, MUSIQ, and MANIQA. FID quantifies the feature-space distribution distance between restored and real-world images. NIQE evaluates naturalness by modeling natural scene statistics. MUSIQ employs a multi-scale Transformer to assess image detail quality across varying resolutions. MANIQA leverages multi-dimensional attention mechanisms to integrate local and global perceptual assessments. Collectively, this metric suite robustly reflects the visual aesthetic performance of our method in complex real-world environments.

4.2 Performance Comparison

We compare GKI-Net against eight state-of-the-art unsupervised dehazing methods [4, 12, 14, 26–28, 31, 32]. Table 1 and Table 2 present the quantitative evaluations across the four selected datasets.

Qualitative Evaluation: Fig. 3 presents visual comparisons on sample images from the NID dataset. GKI-Net demonstrates superior dehazing efficacy. In contrast to Cycle-Dehaze and Diff-Dehazer which exhibit residual haze, or DehazeSB which suffers from overexposure, our method effectively eliminates these artifacts. We achieve an optimal balance between dehazing intensity and image brightness. Regarding overall clarity, DCP, USID-Net, and RefineDNet display excessive contrast. Cycle-Dehaze suffers from severe color shifts and blurring. Conversely, D4, USID-Net, and our method effectively restore feature details and yield clear and visually balanced images. Concerning color fidelity, YOLY, D4, and our method achieve superior performance.

As shown in Fig. 4, we compare visual results on representative images from the SOTS dataset. Both DCP and Cycle-Dehaze suffer from severe color distortion and feature loss, while YOLY struggles to retain details in the sky regions. In terms of overall clarity, USID-Net, D4, and our proposed method achieve the most favorable visual performance. Conversely, RefineDNet and DehazeSB introduce excessive contrast that significantly compromises image brightness. Although Diff-Dehazer yields over-brightened results, its haze removal capability remains inadequate.

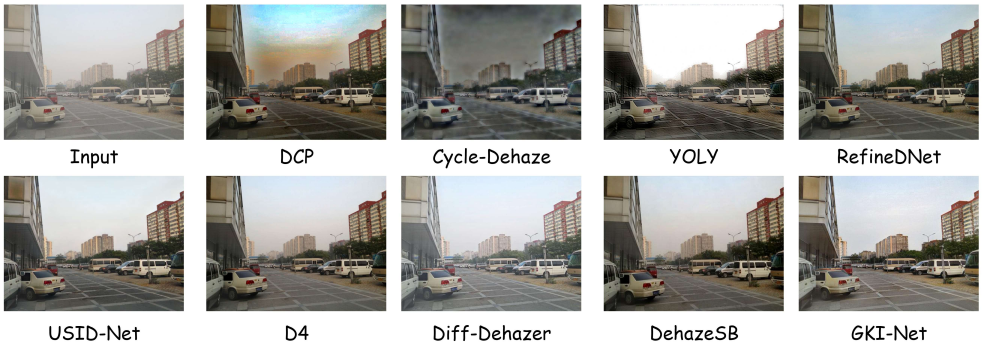


Fig. 4 Visual comparison of GKI-Net with state-of-the-art methods on the SOTS dataset.



Fig. 5 Visual detail comparison of GKI-Net with state-of-the-art methods on the SOTS dataset. Zoomed-in views of the red bounding box are shown for each method.

Detail Restoration Evaluation: To further validate the detail restoration capability of GKI-Net, we select challenging scenes from the SOTS dataset for close-up comparison. Fig. 5 demonstrates that GKI-Net maintains high consistency with state-of-the-art algorithms, such as D4 and USID-Net, in preserving structural details, color saturation, and contrast. Moreover, our method exhibits superior sharpness compared to these baselines. In summary, GKI-Net demonstrates leading performance in both

Table 1 Evaluation results of GKI-Net and SOTA algorithms on SOTS and NID datasets.

Method	SOTS-Indoor		SOTS-Outdoor		NID		Efficiency	
	PSNR↑	SSIM↑	PSNR↑	SSIM↑	PSNR↑	SSIM↑	Params (MB)↓	Flops (G)↓
DCP [4]	13.7	0.71	19.7	0.81	18.9	0.87	×	×
Cycle-Dehaze [12]	21.6	0.92	21.6	0.86	27.3	0.93	11.3	56.9
YOLY [28]	15.8	0.89	15.1	0.86	21.2	0.91	39.9	×
RefineDNet [27]	24.3	0.94	19.8	0.86	29.6	0.91	65.8	75.41
USID-Net [26]	21.5	0.91	18.7	0.84	27.9	0.88	3.83	160
D4 [31]	25.7	0.93	20.9	0.86	28.4	0.92	10.7	2.24
Diff-Dehazer [14]	27.1	0.94	22.7	0.89	31.2	0.96	×	×
DehazeSB [32]	26.3	0.95	21.9	0.91	30.4	0.97	×	×
GKI-Net (Ours)	28.1	0.96	23.8	0.91	31.7	0.98	1.91	23.8

Note: Red: Best performance; Blue: Second-best performance. ↑: Higher is better; ↓: Lower is better.

Table 2 Evaluation results of GKI-Net and SOTA algorithms on RTTS and Haze2020 datasets.

Method	RTTS				Haze2020			
	FID↓	NIQE↓	MUSIQ↑	MANIQA↑	FID↓	NIQE↓	MUSIQ↑	MANIQA↑
DCP [4]	73.1	4.27	52.9	0.12	91.9	3.83	54.4	0.12
Cycle-Dehaze [12]	67.1	4.92	55.7	0.14	81.3	4.23	55.9	0.14
YOLY [28]	70.1	4.65	51.9	0.13	86.5	3.97	54.2	0.14
RefineDNet [27]	65.1	4.01	56.1	0.11	88.2	3.97	54.8	0.12
USID-Net [26]	62.3	4.11	57.2	0.12	86.1	3.98	52.9	0.11
D4 [31]	69.4	4.64	58.4	0.14	85.6	3.97	53.5	0.12
Diff-Dehazer [14]	52.3	3.94	59.2	0.33	71.9	3.57	62.3	0.44
DehazeSB [32]	53.1	3.76	60.1	0.15	69.8	3.74	59.2	0.15
GKI-Net (Ours)	52.8	3.99	60.9	0.14	70.7	3.84	60.4	0.16

Note: Red: Best performance; Blue: Second-best performance. ↑: Higher is better; ↓: Lower is better.

overall visual quality and detail restoration compared to state-of-the-art methods. Furthermore, our approach surpasses existing algorithms in global sharpness.

Quantitative Evaluation: To evaluate the proposed method, we utilize PSNR and SSIM as full-reference metrics while employing FID, NIQE, MUSIQ, and MANIQA for non-reference evaluation across two distinct dataset groups. As reported in Table 1, GKI-Net achieves top-two performance in both PSNR and SSIM on paired datasets. Simultaneously, our model exhibits high efficiency with the lowest parameter count and the second-lowest FLOPs. For the unpaired datasets shown in Table 2, GKI-Net secures the top two positions in FID and MUSIQ scores. It also demonstrates competitive results on NIQE and MANIQA by ranking third and fourth respectively. Given that GKI-Net possesses the fewest parameters among all baselines, these results indicate that our method achieves a superior trade-off between efficiency and performance while maintaining a leading position against state-of-the-art algorithms.

4.3 Ablation Study

To verify the effectiveness and necessity of individual components within GKI-Net, we conduct extensive ablation studies on the real-world NID dataset. We maintain consistent training configurations for all experiments, including learning rate, batch size,

Table 3 Ablation study of GKI-Net core components.

Method	Components			Metrics	
	HPID	ConvKAN	GFSM	PSNR↑	SSIM↑
Baseline	×	×	×	19.4	0.87
Baseline+HPID	✓	×	×	21.7	0.89
Baseline+HPID+ConvKAN	✓	✓	×	27.4	0.94
Baseline+HPID+GFSM	✓	×	✓	30.7	0.96
GKI-Net (Ours)	✓	✓	✓	31.7	0.98

Note: Red: Best; Blue: Second-best. ↑: Higher is better.

Table 4 Comparison of different feature aggregation modules.

Aggregation Module	PSNR↑	SSIM↑
Conv 3×3	27.4	0.94
Non-local Attention [54]	30.3	0.95
Deformable Conv [55]	29.8	0.96
GFSM (Ours)	31.7	0.98

Note: Red: Best; Blue: Second-best. ↑: Higher is better.

and iteration steps, while modifying only the specific module or strategy under verification. Specifically, we perform an in-depth analysis across four dimensions: network architecture design, feature aggregation mechanisms, basis functions of the implicit neural representation, and loss function constraints.

Contribution of Core Components: To isolate the performance gains contributed by individual modules in GKI-Net, we construct a baseline model. This baseline retains the same U-Net backbone as GKI-Net but replaces the GFSM with standard convolutions. Furthermore, we substitute the proposed decoder with a conventional convolutional upsampling module, eliminating both HIPD and ConvKAN to directly regress the dehazed image J. Table 3 demonstrates that all three components play pivotal roles and contribute effective performance gains to GKI-Net. Specifically, our proposed framework achieves the most substantial improvement, surpassing the baseline by 12.3 dB in PSNR and 11 in SSIM on the NID dataset. These results validate the effectiveness of the core architectural design in GKI-Net.

Effectiveness of Gaussian Feature Splatting Module: To verify the superiority of GFSM in feature aggregation, we compare it with representative feature enhancement modules, specifically Deformable Convolution (DCNv2 [55]) and Non-local Attention [54]. The results in Table 4 indicate that GFSM learns sampling offsets and explicitly models spatial influence and importance weights via Gaussian kernel variance and amplitude. This physics-inspired aggregation strategy enables GKI-Net to produce natural transitions in regions with drastic haze density changes, such as sky-building boundaries. Consequently, our module achieves PSNR improvements of 1.9 dB over DCNv2 and 1.4 dB over Non-local Attention.

Table 5 Efficacy of the ConvKAN layer in the HPID.

Layer Type	PSNR↑	SSIM↑
Conv 3×3	27.1	0.91
MLP	29.4	0.95
ConvKAN (Spline=3)	31.7	0.98

Note: Red: Best; Blue: Second-best. ↑: Higher is better.

Table 6 Analysis of multi-scale supervision strategies.

Strategy	Scale			Metrics	
	1/4	1/2	1	PSNR↑	SSIM↑
Single Scale	×	×	✓	27.3	0.92
Dual Scale	✓	×	✓	30.9	0.97
Full Scale	✓	✓	✓	31.7	0.98

Note: Red: Best; Blue: Second-best. ↑: Higher is better.

Impact of ConvKAN in Implicit Representation: GKI-Net leverages ConvKAN layers to process feature transformations for implicit physical parameters. To demonstrate the efficacy of ConvKAN compared to conventional MLPs or CNNs, we conduct comparative experiments. Specifically, we substitute the ConvKAN block in Decoder4, as illustrated in Fig. 2, with standard 3×3 convolutions and MLPs maintaining equivalent channel dimensions. Table 5 reveals that although standard 3×3 convolutions and MLPs require fewer parameters, they are insufficient for modeling the complex nonlinear transmission distribution inherent in the atmospheric scattering model. Conversely, ConvKAN exploits B-spline basis functions to enable high-order nonlinear mappings with minimal parameter cost. Consequently, our module outperforms the standard convolution and MLP baselines, yielding PSNR gains of 4.6 dB and 2.3 dB, respectively.

Multi-Scale Physics Supervision: GKI-Net employs a multi-scale supervision strategy across scales of 1/4, 1/2, and 1. We investigate the impact of different supervision configurations in Table 6. The results indicate that relying solely on the single scale 1 leads to suboptimal performance. This configuration hinders convergence during early training stages and struggles to capture global haze distribution characteristics. In contrast, multi-scale supervision enables the low-resolution branch 1/4 to efficiently estimate the coarse distributions of atmospheric light A and transmission map t, while the high-resolution branch focuses on fine texture restoration. Consequently, the complete three-scale supervision strategy achieves the best quantitative performance.

Impact of Loss Function Terms: To verify the efficacy of the proposed loss function design, we decompose the training objective of GKI-Net into four distinct components: basic reconstruction loss $\mathcal{L}_{rec}$, physical consistency loss $\mathcal{L}_{phy}$, intermediate variable regularization $(\mathcal{L}_{tv} + \mathcal{L}_{smooth_A})$, and visual prior constraints $(\mathcal{L}_{edge} + \mathcal{L}_{dc})$. We detail

Table 7 Ablation study on different loss components.

$\mathcal{L}_{\text{rec}}$	$\mathcal{L}_{\text{phy}}$	$\mathcal{L}_{\text{tv}}+\mathcal{L}_{\text{smooth_A}}$	$\mathcal{L}_{\text{edge}}+\mathcal{L}_{\text{dc}}$	PSNR↑	SSIM↑
✓	×	✓	✓	29.7	0.87
✓	✓	×	✓	31.4	0.97
✓	✓	✓	×	30.9	0.96
✓	✓	✓	✓	31.7	0.98

Note: Red: Best; Blue: Second-best. ↑: Higher is better.

the ablation study regarding these key terms in Table 7. Quantitative results indicate that each component provides essential supervision, and the joint optimization of all loss functions yields optimal performance.

5 Conclusion

We propose GKI-Net, a physics-embedded dehazing framework that integrates Gaussian splatting with Convolutional Kolmogorov-Arnold Networks. By utilizing a Gaussian Feature Splatting Module to adaptively aggregate global context and leveraging the efficient nonlinear transformation capabilities of ConvKAN, our network explicitly decouples and precisely reconstructs the transmission map and atmospheric light defined in the atmospheric scattering model. Extensive experiments on four benchmark datasets demonstrate that GKI-Net significantly outperforms state-of-the-art methods in both quantitative metrics and visual fidelity. Furthermore, our method guarantees the physical consistency of restored results and substantially enhances the environmental adaptability and perceptual robustness of visual sensing systems under adverse weather conditions.

Acknowledgements. The authors thank Google Gemini, Google Translate, and Grammarly for their assistance in language translation, editing, and grammar correction during the preparation of this manuscript.

Declarations

- **Funding**: This work was supported by the Jiaxing Public Welfare Research Project under grant no. 2025CGW075, the Jiaxing Team Science and Technology Commissioner Project under grant no. K2024-1-1-06, and the General Scientific Research Project of Zhejiang Provincial Department of Education under grant no. Y202456753.
- **Competing interests**: The authors declare no conflicts of interest.
- **Code availability**: Not applicable.
- **Authors' contributions**:Conceptualization, x.w.; methodology, Z.L.; software, x.w.; validation, Z.L. and B.L.; formal analysis, B.L.; investigation, x.w., Y.C.; resources, X.W.; data curation, B.L.; writing—original draft preparation, X.W.; writing—review and editing, B.L. and Z.L.; visualization, Z.L., Y.C.; supervision, Z.L., Y.C.; project administration, X.W.; funding acquisition, X.W., Y.C.. All authors have read and agreed to the published version of the manuscript.

References

[1] Ayoub, A., El-Shafai, W., El-Samie, F.E.A., *et al.*: Review of dehazing techniques: Challenges and future trends. Multimedia Tools Appl. **84**(3), 1103–1131 (2025)

[2] Agrawal, S.C., Jalal, A.S.: A comprehensive review on analysis and implementation of recent image dehazing methods. Arch. Comput. Methods Eng. **29**(7), 4799–4850 (2022)

[3] Li, C., Yan, W., Zhao, H., *et al.*: Tffd-net: An effective two-stage mixed feature fusion and detail recovery dehazing network. Vis. Comput. **41**(6), 4001–4016 (2025)

[4] He, K., Sun, J., Tang, X.: Single image haze removal using dark channel prior. IEEE Trans. Pattern Anal. Mach. Intell. **33**(12), 2341–2353 (2010)

[5] Zhu, Q., Mai, J., Shao, L.: A fast single image haze removal algorithm using color attenuation prior. IEEE Trans. Image Process. **24**(11), 3522–3533 (2015)

[6] Cai, B., Xu, X., Jia, K., *et al.*: Dehazenet: An end-to-end system for single image haze removal. IEEE Trans. Image Process. **25**(11), 5187–5198 (2016)

[7] Li, B., Peng, X., Wang, Z., *et al.*: Aod-net: All-in-one dehazing network. In: Proc. IEEE Int. Conf. Comput. Vis. (ICCV), pp. 4770–4778 (2017)

[8] Ren, W., Liu, S., Zhang, H., *et al.*: Single image dehazing via multi-scale convolutional neural networks. In: Proc. Eur. Conf. Comput. Vis. (ECCV), pp. 154–169. Springer, Cham (2016)

[9] Qin, X., Wang, Z., Bai, Y., *et al.*: Ffa-net: Feature fusion attention network for single image dehazing. In: Proc. AAAI Conf. Artif. Intell., vol. 34, pp. 11908–11915 (2020)

[10] Liu, X., Ma, Y., Shi, Z., *et al.*: Griddehazenet: Attention-based multi-scale network for image dehazing. In: Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV), pp. 7314–7323 (2019)

[11] Deng, Q., Huang, Z., Tsai, C.C., *et al.*: Hardgan: A haze-aware representation distillation gan for single image dehazing. In: Proc. Eur. Conf. Comput. Vis. (ECCV), pp. 722–738. Springer, Cham (2020)

[12] Engin, D., Genç, A., Ekenel, H.K.: Cycle-dehaze: Enhanced cyclegan for single image dehazing. In: Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW), pp. 825–833 (2018)

[13] Chen, Z., Wang, Y., Yang, Y., *et al.*: Psd: Principled synthetic-to-real dehazing guided by physical priors. In: Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR), pp. 7180–7189 (2021)

[14] Lan, Y., Cui, Z., Liu, C., *et al.*: Exploiting diffusion prior for real-world image dehazing with unpaired training. In: Proc. AAAI Conf. Artif. Intell., vol. 39, pp. 4455–4463 (2025)

[15] Liu, C., Qi, L., Pan, J., et al.: Frequency domain-based diffusion model for unpaired image dehazing. arXiv:2507.01275 (2025)

[16] Kerbl, B., Kopanas, G., Leimkühler, T., Drettakis, G.: 3d gaussian splatting for real-time radiance field rendering. ACM Trans. Graph. **42**(4), 139–113914 (2023)

[17] Chen, Y., Yu, W., Li, G., et al.: LL-GaussianImage: Efficient image representation for zero-shot low-light enhancement with 2D Gaussian splatting. arXiv:2601.15772 (2026)

[18] Zhang, X., Ge, X., Xu, T., *et al.*: Gaussianimage: 1000 fps image representation and compression by 2d gaussian splatting. In: Proc. Eur. Conf. Comput. Vis. (ECCV), pp. 327–345. Springer, Cham (2024)

[19] Ma, C., Zhao, J., Chen, J.: Dehazegs: 3d gaussian splatting for multi-image haze removal. IEEE Signal Process. Lett. (2025)

[20] Yu, J., Wang, Y., Jiang, A., et al.: DehazeGS: Seeing through fog with 3D Gaussian splatting. arXiv:2501.03659 (2025)

[21] Xu, J., Wang, H., Li, M., et al.: Dehazesplat: Unifying unsupervised dehazing and 3d gaussian splatting for robust reconstruction in hazy scenes. SSRN Electron. J. (2025)

[22] Jain, N., Jong, A., Scherer, S., et al.: SmokeSeer: 3D Gaussian splatting for smoke removal and scene reconstruction. arXiv:2509.17329 (2025)

[23] Qu, Y., Chen, Y., Huang, J., *et al.*: Enhanced pix2pix dehazing network. In: Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR), pp. 8160–8168 (2019)

[24] Zhu, J.Y., Park, T., Isola, P., *et al.*: Unpaired image-to-image translation using cycle-consistent adversarial networks. In: Proc. IEEE Int. Conf. Comput. Vis. (ICCV), pp. 2223–2232 (2017)

[25] Yang, X., Xu, Z., Luo, J.: Towards perceptual image dehazing by physics-based disentanglement and adversarial training. In: Proc. AAAI Conf. Artif. Intell., vol. 32 (2018)

[26] Li, J., Li, Y., Zhuo, L., *et al.*: Usid-net: Unsupervised single image dehazing network via disentangled representations. IEEE Trans. Multimedia **25**, 3587–3601 (2022)

[27] Zhao, S., Zhang, L., Shen, Y., *et al.*: Refinednet: A weakly supervised refinement

framework for single image dehazing. IEEE Trans. Image Process. **30**, 3391–3404 (2021)

[28] Li, B., Gou, Y., Gu, S., *et al.*: You only look yourself: Unsupervised and untrained single image dehazing neural network. Int. J. Comput. Vis. **129**(5), 1754–1767 (2021)

[29] Park, T., Efros, A.A., Zhang, R., *et al.*: Contrastive learning for unpaired image-to-image translation. In: Proc. Eur. Conf. Comput. Vis. (ECCV), pp. 319–345. Springer, Cham (2020)

[30] Wang, Y., Yan, X., Wang, F.L., *et al.*: Ucl-dehaze: Toward real-world image dehazing via unsupervised contrastive learning. IEEE Trans. Image Process. **33**, 1361–1374 (2024)

[31] Yang, Y., Wang, C., Liu, R., *et al.*: Self-augmented unpaired image dehazing via density and depth decomposition. In: Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR), pp. 2037–2046 (2022)

[32] Lan, Y., Cui, Z., Luo, X., *et al.*: When schrodinger bridge meets real-world image dehazing with unpaired training. In: Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV), pp. 8756–8765 (2025)

[33] Huang, B., Yu, Z., Chen, A., *et al.*: 2d gaussian splatting for geometrically accurate radiance fields. In: Proc. ACM SIGGRAPH Conf., pp. 1–11 (2024)

[34] Hanson, A., Tu, A., Lin, G., *et al.*: Speedy-splat: Fast 3d gaussian splatting with sparse pixels and sparse primitives. In: Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR), pp. 21537–21546 (2025)

[35] Yan, Y., Lin, H., Zhou, C., *et al.*: Street gaussians: Modeling dynamic urban scenes with gaussian splatting. In: Proc. Eur. Conf. Comput. Vis. (ECCV), pp. 156–173. Springer, Cham (2024)

[36] Zhou, X., Lin, Z., Shan, X., *et al.*: Drivinggaussian: Composite gaussian splatting for surrounding dynamic autonomous driving scenes. In: Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR), pp. 21634–21643 (2024)

[37] Cheng, K., Long, X., Yang, K., *et al.*: Gaussianpro: 3d gaussian splatting with progressive propagation. In: Proc. 41st Int. Conf. Mach. Learn. (ICML) (2024)

[38] Liu, Y., Luo, C., Fan, L., *et al.*: Citygaussian: Real-time high-quality large-scale scene rendering with gaussians. In: Proc. Eur. Conf. Comput. Vis. (ECCV), pp. 265–282. Springer, Cham (2024)

[39] Fan, J., Li, W., Han, Y., *et al.*: Momentum-gs: Momentum gaussian self-distillation for high-quality large scene reconstruction. In: Proc. IEEE/CVF Int.

Conf. Comput. Vis. (ICCV), pp. 25250–25260 (2025)

[40] Xu, Y., Zhang, J., Chen, Y., et al.: PMGS: Reconstruction of projectile motion across large spatiotemporal spans via 3D Gaussian splatting. arXiv:2508.02660 (2025)

[41] Xu, Y., Zhang, J., Liu, H., et al.: PEGS: Physics-event enhanced large spatiotemporal motion reconstruction via 3D Gaussian splatting. arXiv:2511.17116 (2025)

[42] Yi, T., Fang, J., Wang, J., *et al.*: Gaussiandreamer: Fast generation from text to 3d gaussians by bridging 2d and 3d diffusion models. In: Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR), pp. 6796–6807 (2024)

[43] Ni, C., Zhao, G., Wang, X., *et al.*: Recondreamer: Crafting world models for driving scene reconstruction via online restoration. In: Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR), pp. 1559–1569 (2025)

[44] Zhu, L., Lin, G., Chen, J., *et al.*: Large images are gaussians: High-quality large image representation with levels of 2d gaussian splatting. In: Proc. AAAI Conf. Artif. Intell., pp. 10977–10985 (2025)

[45] Zeng, Z., Wang, Y., Guan, T., *et al.*: Instant gaussianimage: A generalizable and self-adaptive image representation via 2d gaussian splatting. In: Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV), pp. 27896–27905 (2025)

[46] Jiang, C., Li, Z., Zhao, H., et al.: Beyond pixels: Efficient dataset distillation via sparse Gaussian representation. arXiv:2509.26219 (2025)

[47] Omri, Y., Ding, C., Weissman, T., Tambe, T.: Vision-language alignment from compressed image representations using 2D Gaussian splatting. arXiv:2509.22615 (2025)

[48] Hu, J., Xia, B., Chen, B., *et al.*: Gaussiansr: High fidelity 2d gaussian splatting for arbitrary-scale image super-resolution. In: Proc. AAAI Conf. Artif. Intell., vol. 39, pp. 3554–3562 (2025)

[49] Chen, Y., Fang, Y., Li, G., et al.: LL-GaussianMap: Zero-shot low-light image enhancement via 2D Gaussian splatting guided gain maps. arXiv:2601.15766 (2026)

[50] Li, B., Ren, W., Fu, D., *et al.*: Benchmarking single-image dehazing and beyond. IEEE Trans. Image Process. **28**(1), 492–505 (2018)

[51] Chi, K., Yuan, Y., Wang, Q.: Trinity-net: Gradient-guided swin transformer-based remote sensing image dehazing and beyond. IEEE Trans. Geosci. Remote Sens. **61**, 1–14 (2023)

[52] Ancuti, C.O., Ancuti, C., Timofte, R.: Nh-haze: An image dehazing benchmark with non-homogeneous hazy and haze-free images. In: Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW), pp. 444–445 (2020)

[53] Shao, Y., Li, L., Ren, W., *et al.*: Domain adaptation for image dehazing. In: Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR), pp. 2808–2817 (2020)

[54] Zhang, Y., Li, K., Li, K., et al.: Residual non-local attention networks for image restoration. arXiv:1903.10082 (2019)

[55] Zhu, X., Hu, H., Lin, S., *et al.*: Deformable convnets v2: More deformable, better results. In: Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR), pp. 9308–9316 (2019)