Gated Graph Reasoning Network with Color Tuning and Calibration for Image Dehazing

Jiajia Hu[1], Wujin Li[1], Huifang Liang[2,*]

[1]Information Science and Technology Laboratory, Hunan Institute of Science and Technology, Yueyang 414006, China

[2]School of Business, Guilin University of Electronic Technology, Guilin 541004, China

*Corresponding email: huifangliang@163.com

Abstract

Removing non-homogeneous haze from a single image without losing structural and color fidelity is a major challenge due to the lack of natural scene priors. To address this problem, we propose a novel Gated Graph Reasoning Network with Color Tuning and Calibration (GCTC-Net). Specifically, to eliminate the information redundancy of traditional rigid feature stacking, an Adaptive Gated Attention Interaction (AGAI) module is introduced to adaptively aggregate multi-exposure Gated Contrast-enhanced Prior (GCP) maps with primary hazy features via pixel-level attention. Subsequently, a dual-path topo-logical graph reasoning backbone is deployed to model non-local dependencies across spatial and channel dimensions, effectively propagating clean structural clues to heavily contaminated nodes. Finally, a dedicated non-linear Color Calibration Module (CCM) is incorporated at the back end to execute localized color variance restoration, successfully suppressing cumulative chromatic deviations and white-balance shifts. Extensive experiments on public benchmarks demonstrate that our proposed GCTC-Net achieves superior performance and establishes an optimal balance between structural sharpness and realistic color fidelity.

Keywords

Non-homogeneous haze removal, Gated prior fusion, Graph reasoning, Color calibration

Introduction

In real-world scenarios, atmospheric scattering inherently exhibits severe spatial non-homogeneity, where uneven haze densities localized across distinct regions obscure critical high-frequency structural elements and introduce complex degradation patterns. Recently, pioneering methodologies have attempted to address these spatial variations by incorporating artificial multi-exposure scene priors and leveraging Graph Convolutional Networks (GCNs) to capture global context and long-range dependencies. Despite their commendable performance gains, two fundamental structural bottlenecks continue to impede their practical efficacy in demanding environmental restoration tasks.

On the one hand, existing multi-exposure networks typically rely on a primitive and rigid channel concatenation strategy to aggregate diverse artificial priors. This brute-force feature stacking operates under a uniform assumption that fails to adapt to spatial density fluctuations, inevitably introducing massive information redundancy. Consequently, such excessive multi-exposure data inadvertently overwhelms and corrupts the latent primary feature representations across uneven varying haze densities. To resolve these cross-scale variances and extract clean spatial distributions, contemporary deep learning frameworks have explored parallel multi-scale attention mapping paradigms to distill vital localized structural textual cues [1].

On the other hand, executing extensive topological propagation within deep graph reasoning frameworks acts as a double-edged sword. While global non-local feature aggregation successfully reconstructs long-range pixel dependencies, it inherently dilutes sensitive channel-wise variations and spatial local boundaries. This unbound cross-node feature blending frequently triggers cumulative chromatic aberrations and implicit white-balance shifts. Without the constraint of an explicit, cross-domain color regularizer, these graph-based frameworks systematically propagate errors throughout the latent space, ultimately producing unnatural color oversaturation and severe color casting in the final restored outputs.

To resolve these interconnected bottlenecks, we propose a novel Gated Graph Reasoning Network with Color Tuning and Calibration (GCTC-Net) for robust single image haze removal. Rather than expanding upon traditional parameter estimation pipelines, our structural backbone embraces advanced multi-scale feature fusion network paradigms that exploit explicit feature differences to dynamically handle non-uniform atmospheric densities [2]. Furthermore, to effectively prevent the erosion of delicate edge profiles and subtle fine boundaries during global graph reasoning propagation, our architecture integrates localized spatial partitioning concepts inspired by mainstream deep learning-based image segmentation frameworks [3]. Finally, to address the severe ambient light attenuation and color degradation caused by dense aerosol scattering, we formulate a bounded regularizer drawing inspiration from the domain-adaptive illumination adjustments of recent low-light image enhancement networks [4]. By integrating these components into a unified pipeline, the proposed GCTC-Net explicitly suppresses topological-induced color discrepancies, delivering exceptional visual naturalness and color fidelity.

The primary academic contributions of this work are summarized as follows: (1) Holistic framework realization: This paper introduces GCTC-Net, a comprehensive end-to-end framework that explicitly decouples structural reconstruction from color appearance adjustment. This dual-path optimization architecture uniquely balances high-frequency structural crispness and global aesthetic color fidelity under highly non-homogeneous haze distributions. (2) Adaptive gated interaction: This paper designs an AGAI module as a dynamic front-end filtering interface. Governed by a pixel-level attention mechanism, the AGAI module selectively routes and aggregates salient multi-exposure priors while blocking redundant feature components, effectively resolving the performance degradation inherent in traditional rigid prior stacking. (3) Advanced color calibration: This paper formulates a dedicated non-linear Color Calibration Module acting as a back-end post-processing engine. By implementing a localized color variance calibration network with bounded constraints, the CCM explicitly suppresses the cumulative white-balance shifts and chromatic distortions induced by deep graph propagation, guaranteeing exceptional perceptual realism.

Related work

Early single image dehazing approaches relied on physical statistics and handcrafted prior assumptions to recover scene visibility. Traditional algorithms like contrast limited adaptive histogram equalization (CLAHE) were deployed to enhance local dynamic ranges. To separate texture components from complex backgrounds, the discrete wavelet transforms de-composed structural details into sub-bands, while the center/surround Retinex theory estimated illumination tracks [5,6]. A major milestone is the Dark Channel Prior (DCP), which estimates transmission maps based on low-intensity pixels in non-sky patches. Following this path, the color attenuation prior and multiple priors constraints were proposed to solve the ill-posed atmospheric scattering model [7-9]. However, these prior-driven methods remain prone to severe color distortion and halo artifacts when empirical statistical assumptions fail in real-world scenarios.

With the advancement of deep learning, Convolutional Neural Networks (CNNs) shifted the paradigm by learning direct mapping from hazy inputs to clear counterparts. To bypass the reliance on paired training data, recent frameworks explore unsupervised image dehazing architectures [10]. To preserve spectral properties, frequency-based attention within Generative Adversarial Networks (GANs) balances high-frequency components, while advanced spatial frequency attention networks like T-SFANet enhance non-local interactions [11,12]. For parameter estimation, DehazeNet established an early system to estimate media transmission maps, paving the way for comprehensive end-to-end networks that restore images directly without physical parameter reliance [13]. To break the performance plateaus of early convolutional systems, modern deep architectures leverage structural operations such as dual residual networks to amplify the latent potential of paired restoration blocks, or employ attention-based multi-scale
grid network configurations to enhance information flow across diverse resolutions [14,15]. Despite their structural capabilities, standard CNNs remain

fundamentally constrained by local receptive fields, failing to model long-range spatial dependencies cleanly.

Graph Convolutional Networks in image restoration

To break the spatial limitations of local convolutions, Graph Convolutional Networks (GCNs) have been successfully translated from non-Euclidean data to grid-structured computer vision tasks [16]. By treating pixels, patches, or feature channels as topological nodes, graph-based methods construct non-local relational graphs to propagate complementary information across long spatial distances. This global reasoning capability makes GCNs exceptionally powerful for capturing large-scale structural context and global degradation distributions. Nevertheless, deploying graph networks for high-fidelity image restoration remains a double-edged sword. While deep non-local graph propagation successfully reconstructs long-range structural dependencies, the global cross-node feature aggregation often dilutes sensitive local color variations and cross-domain boundary distributions. Without localized guidance, extensive graph propagation inevitably induces a white-balance shift and cumulative color distortion, frequently manifesting as unnatural saturation or severe color cast in final predictions.

Color preservation in dehazing frameworks

Retaining aesthetic color realism and preventing chromatic deviation is a crucial criterion for high-quality image dehazing. Although many advanced models achieve notable scores in structural evaluation metrics, they often compromise the visual fidelity of background textures, generating dark artifacts or severe color bias. This phenomenon is particularly aggravated in non-homogeneous outdoor environments, where severe light attenuation and dense scattering corrupt the inherent ambient illumination.

To combat these illumination and color inconsistencies, contemporary learning frameworks have introduced diverse architectural regularizers. These include generative adversarial paradigms equipped with fusion-discriminators to enforce global fidelity. They also include multi-scale boosted architecture utilizing dense feature fusion to maintain spectral stability across progressive depths [17,18].

To safeguard realistic color balance, recent frameworks have explored diverse color-aware constraints or downstream restoration layers. However, existing schemes primarily embed these constraints within the training loss functions, which provides implicit regularization but fails to offer an explicit topological solution during the inference phase. Driven by this insight, structural reconstruction and visual appearance adjustment are deliberately decoupled in the proposed architecture. By combining an adaptive gated attention module (GF_layer) at the front-end and an independent non-linear CCM at the back end, the proposed model effectively rectifies color discrepancies while fully preserving the non-local topo-logical representations generated by the graph backbone.

Proposed method

Overall pipeline

The comprehensive architecture of the proposed Gated Graph Reasoning Network with Color Tuning and Calibration (GCTC-Net) is illustrated in Figure 1. Given a single non-homogeneous hazy image $I \in \mathbf{R}^{H\times W \times C}$, this paper framework aims to reconstruct a high-fidelity, haze-free image I_{out} while simultaneously preserving structural sharpness and realistic color fidelity. Historically, addressing this ill-posed inverse problem involved analyzing the inherent physics of atmospheric propagation, such as utilizing polarization-based vision methods to decouple the air light component or designing blind haze separation paradigms based on physical statistics. While mathematically elegant, these classic methods struggle under severely non-homogeneous concentrations [19].

To bypass these limitations while retaining physical awareness, the holistic forward propagation of GCTC-Net is structurally decoupled into three distinct stages. First, a twelve-channel GCP maps is generated via dynamic parameter estimation to serve as a multi-exposure contrast navigator, drawing inspiration from classical physical contrast restoration concepts under varying weather conditions. Second, to resolve the severe information redundancy inherent in rigid feature stacking, the primary convolutional features and the GCP map is transmitted into the AGAI module to adaptively filter out irrelevant degradation noise. Third, the purified representation is forwarded to a bidimensional topological graph reasoning backbone to explore non-local dependencies across spatial and channel domains, followed by a dedicated non-linear CCM at the back end to explicitly rectify cumulative color distortions and balance shifts.

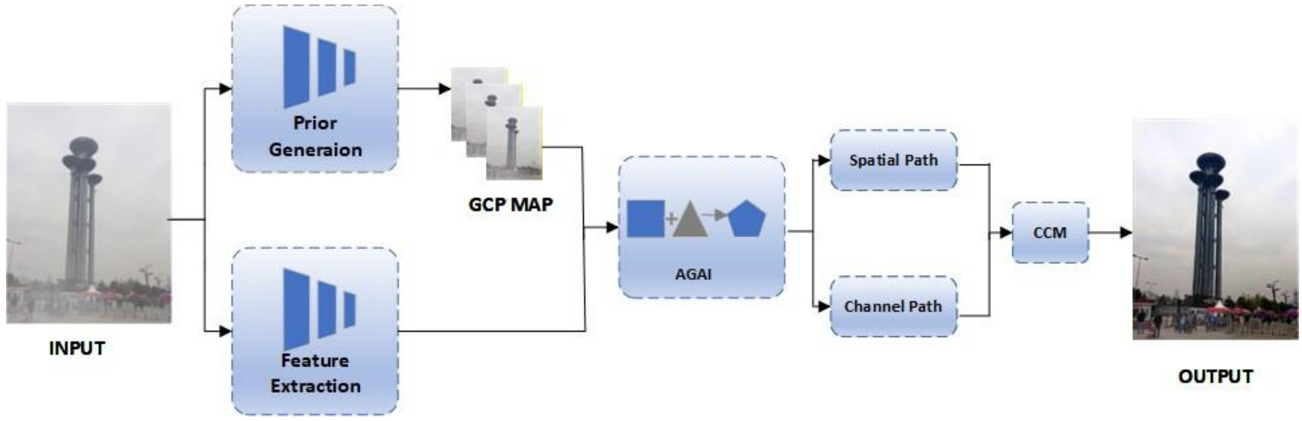

Figure 1. The comprehensive architecture of the proposed GCTC-Net.

Adaptive gated prior-feature interaction

Standard Red-Green-Blue (RGB) representations lack sufficient sensitivity to indicate non-uniform degradation variations across disparate spatial regions. To address this, we utilize a lightweight parameter estimator to predict two dynamic scene-specific scaling parameters α and β from the input hazy profile. Four distinct non-linear exponential mapping operations Φ_k $(\cdot)$ are executed simultaneously to simulate virtual multi-exposure shots P_k ($k \in \{1, 2, 3, 4\}$), which act as pseudo-priors to bound the atmospheric luminance attenuation:

$$P_k(x) = \Phi_k(I(x), \alpha, \beta) \tag{1}$$

Stacking these four generated virtual exposure shots along the channel dimension yields the comprehensive twelve-channel GCP Map $M_{GCP} \in \mathrm{R}^{H\times W\times 12}$.

To bypass noise corruption and performance degradation stemming from rigid channel stacking, the primary feature map $F_{hazy} \in \mathrm{R}^{H\times W\times C}$ and the pseudo-prior M_{GCP} are injected into the AGAI module (GF_layer). Instead of treating all stacked channels uniformly, the AGAI module computes pixel-level spatial gating weights $Y_1, Y_2 \in \mathrm{R}^{H\times W\times 1}$ to modulate the latent information flow. It derives the purified hybrid feature

map $F_{fusion} \in \mathrm{R}^{H\times W\times C}$ via element-wise gating:

$$Y_1 = \sigma\left(\mathrm{Conv}_{1\times 1}([F_{hazy}, M_{GCP}])\right), Y_2 = 1 - Y_1 \tag{2}$$

$$F_{fusion} = Y_1 \odot \mathrm{Conv}(F_{hazy}) + Y_2 \odot \mathrm{Conv}(M_{GCP}) \tag{3}$$

where $[\cdot,\cdot]$ denotes channel-wise concatenation, σ represents the standard Sigmoid activation function, and $\odot$ denotes the Hadamard product. This explicit gate modulation allows the network to adaptively route illumination-enhanced components while suppressing redundant background clutter, and the visualized effect is shown in Figure 2.

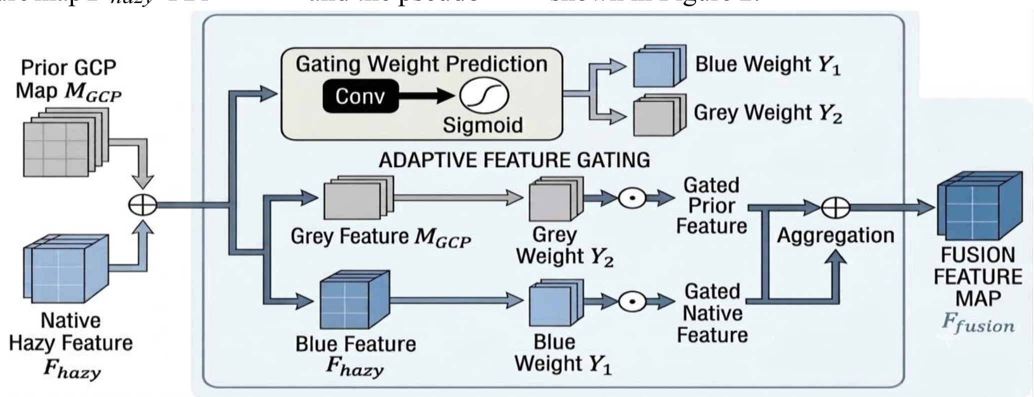

Figure 2. The internal structure of AGAI.

Bidimensional graph reasoning and color calibration

The purified hybrid features F_{fusion} are subsequently processed by a bidimensional topological graph network to capture long-range contextual relationships, followed by a post-calibration mechanism. Within non-homogeneous scattering environments, structural attenuation and visibility degradation are heavily tied to spatial variations and depth-dependent medium characteristics.

To effectively model these uneven distributions without the receptive field constraints of standard convolutions, our architecture mapping decouples global graph reasoning into concurrent spatial and channel domains, drawing structural inspiration from non-local affinity propagation paradigms traditionally optimized for image restoration [20].

(1) Spatial-Topological Graph Reasoning (STGR): Features are projected into spatial node embeddings $X_s \in \mathrm{R}^{N\times C}$ ($N = H' \times W'$), where each node encapsulates a localized patch identity. The non-local spatial adjacency matrix $A_s \in \mathrm{R}^{N\times N}$, which computes the pixel-to-pixel affinity mapping across distant coordinates, and the propagated representations Zs are defined as:

$$A_s = \mathrm{Softmax}(\theta(X_s)\phi(X_s)^T),\ Z_S = \mathrm{ReLU}(A_S X_S W_S) \quad (4)$$

where $\theta(\cdot)$ and $\phi(\cdot)$ denote embedding functions that project spatial relationships into a low-dimensional manifold, and Ws represents a learnable weight matrix regulating graph message passing.

(2) Cross-Channel Graph Reasoning (CCGR): Concurrently, to explore cross-domain correlations and spectral dependencies that are often corrupted under varying aerosol densities, representations are transposed to channel node embeddings $X_c \in \mathrm{R}^{C\times N}$.

The channel adjacency matrix $A_c \in \mathrm{R}^{C\times C}$, which explicitly models global relational properties among distinct feature channels, and the reasoned features Z_c are formulated as:

$$A_c = \mathrm{Softmax}(\psi(X_c)\omega(X_c)^T), Z_c = \mathrm{ReLU}(A_c X_c W_c) \quad (5)$$

where $\psi(\cdot)$ and $\omega(\cdot)$ are linear transformation operators, and W_c parameterizes the inter-channel feature transition. The spatial and channel outputs are aggregated and reshaped back to the coordinate grid domain to generate the unified topological representation $F_{graph} \in \mathrm{R}^{H\times W\times C}$. Finally, to suppress the white-balance shifts and cumulative color distortions systematically induced by such extensive deep graph propagation, the non-linear CCM (i.e., CVR_Layer) executes a bounded residual adjustment to output the final clean feature map F_{final}:

$$\Delta F_{color} = \mathrm{Tanh}\left(\mathrm{Conv}_{1\times1}(F_{graph})\right) \quad (6)$$

$$F_{final} = F_{graph} + \gamma \cdot \Delta F_{color} \quad (7)$$

where the non-linear Tanh $(\cdot)$ function restricts the variance range within a bounded stable zone, preventing unnatural oversaturation, and γ=0.1 is an empirical scaling factor ensuring stable color calibration, and the illustrated effect is shown in Figure 3.

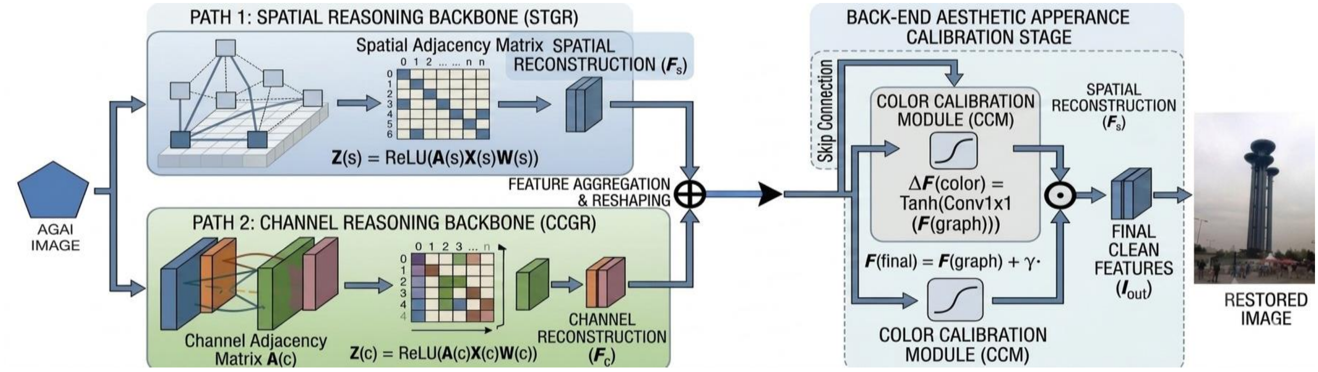

Figure 3. Detailed architecture of the dual-path topological graph reasoning and color calibration stage. It consists of STGR, CCGR, and the back-end non-linear CCM.

Experiments and analysis

Experimental setup and dataset protocols

To ensure robust generalization capability and fairly evaluate GCTC-Net against contemporary state-of-the-art (SOTA) frameworks, we establish a rigorous cross-dataset experimental protocol. Specifically, our model is trained exclusively in large-scale synthetic benchmarks and subsequently evaluated on diverse, highly challenging real-world hazy distributions without any fine-tuning.

(1) Training configuration: The training phase is carried out utilizing the widely acknowledged Realistic Single Image Dehazing (RESIDE) dataset, leveraging both its indoor and outdoor synthesis engines to provide a diverse spectrum of scattering distributions:

RESIDE Indoor Training Set (ITS): This set comprises 13,990 pairs of synthetic indoor hazy images generated from clean ground-truth profiles via the physical atmospheric scattering model. It provides dense structural variations and localized texture boundaries, forcing the network to learn basic structural reconstruction paradigms.

RESIDE Outdoor Training Set (OTS) -β: To expose the framework to expansive outdoor illumination and diverse depth fields, we incorporate the OTS-β partition, which contains 72,135 pairs of outdoor hazy scenes. Utilizing this massive synthetic corpus allows our bi-dimensional graph backbone to adequately regularize its topological relational matrices (A_s and A_c) prior to evaluating real-world degradation.

(2) Testing configuration: To verify the out-of-distribution generalization and practical efficacy of GCTC-Net in handling complex atmospheric scattering, we deploy three distinct real-world benchmarks as our testbed, completely unseen during the training process.

a) Non-Homogeneous Haze (NH-HAZE) dataset: Serving as the primary crucible for non-homogeneous evaluation, this benchmark provides real out-door scenes where smoke generators introduce highly uneven, cloud-like hazy distributions. It evaluates the network's upper-bound capacity to filter feature redundancy via the AGAI module.

b) Outdoor Haze (O-HAZE) dataset: This dataset consists of real outdoor scenes captured under controlled uniform haze. It provides an ideal platform to test the model's structural preservation and global ambient light estimation when confronting real-world aerosol scattering.

c) Dehazing Quality (DHQ) dataset: To assess the perceptual naturalness and real-world applicability of our framework, we further introduce the DHQ benchmark. Comprising diverse real-world hazy images captured in wild environments without paired ground-truth, the DHQ dataset evaluates whether our back-end non-linear CCM can successfully suppress white-balanced shifts and chromatic aberrations under arbitrary, unconstrained illumination tracks.

Experimental setup and details

Datasets: The proposed model is evaluated on two widely used public benchmarks: O-HAZE for real homogeneous haze, and NH-HAZE for real non-homogeneous haze. Specifically, NH-HAZE exhibits unevenly distributed and highly dense cloud-like haze, which serves as our primary benchmark to validate the model's capability under non-uniform degradation.

Evaluation metrics: To comprehensively evaluate structural restoration and perceptual color fidelity, we employ two standard metrics. Peak Signal-to-Noise Ratio (PSNR), Structural Similarity Index Measure (SSIM), and Learned Perceptual Image Patch Similarity (LPIPS). A lower LPIPS score indicates higher structural realism and better perceptual consistency with the ground-truth images.

Implementation details: The proposed GCTC-Net is implemented in PyTorch and trained on a computing platform equipped with an Intel Core i7 3.60 GHz processor and an NVIDIA RTX 4060Ti 16GB graphics processing unit (GPU). The training images are randomly cropped to 256 × 256 patches. We utilize the Adam optimizer with β_1=0.900 and β_2=0.999. The initial learning rate is set to 2×10^{-4} and steadily decayed to 1×10^{-6} using a cosine annealing strategy.

Quantitative analysis with IQA metrics

To comprehensively evaluate the perceptual quality of the restored images, we conduct a dedicated evaluation leveraging advanced image quality assessment (IQA) metrics. This analysis explicitly decouples the performance evaluation into two structural dimensions: Full-reference IQA (FQ-IQA) for paired non-homogeneous benchmarks, and no-reference IQA (NR-IQA) for unconstrained, open-world scenes. This dual-track protocol serves as a rigorous benchmark against contemporary representative pipelines, which encompass advanced Taylor-expansion Transformers, transmission-aware 3D position attention network architectures, and hybridized remote sensing residual attention fusions [21]. Concurrently, as systematically explored in standard benchmarking frameworks like RESIDE, relying solely on conservative, pixel-wise distortion constraints (e.g., PSNR and SSIM) often fails to accurately reflect human visual preferences and high-level scene understanding under severe semantic foggy degradation [22,23]. The holistic quantitative evaluations across these IQA paradigms are systematically reported in Table 1 and Table 2.

Table 1. Comparison of traditional dehazing algorithms.

Method		DCP	CAP	Meng	AMEF	DehazeNet	AOD-Net	GFN
FQ IQA	NH-HAZE	56.92	32.79	50.32	42.04	29.02	24.72	27.65
	O-HAZE	54.10	53.52	53.32	69.67	53.84	59.35	63.28

Method		DCP	CAP	Meng	AMEF	DehazeNet	AOD-Net	GFN
NR IQA	DHQ	48.04	45.77	41.09	48.78	48.27	48.48	49.30

Table 2. Additional ablation results on real-world datasets.

Method		DCPDN	GCANet	DuRN	GridNet	EDN	MCN	GCTC-Net (Ours)
FQ IQA	NH-HAZE	34.68	46.24	74.86	76.59	72.76	76.22	77.31
	O-HAZE	51.92	61.99	81.97	80.69	81.66	77.82	77.82
NR IQA	DHQ	48.02	47.35	48.29	49.09	49.29	48.93	49.55

On the demanding non-homogeneous benchmarks (NH-HAZE and O-HAZE), GCTC-Net delivers a highly competitive performance margin under the FQ-IQA protocol. Compared to uniform scattering profiles, real-world non-homogeneous dense haze introduces severe spatial anisotropy and non-uniform density variations. Under this challenging setup, on the NH-HAZE dataset, our model secures the optimal score of 77.31, suppressing the runner-up GridNet (76.59) and the close baseline MCN (76.22), while significantly outperforming traditional multi-scale boosted dense fusion convolutional frameworks. On the uniform O-HAZE dataset, although DuRN maintains the peak position with a score of 81.97, GCTC-Net achieves a robust performance of 77.82, performing on par with MCN (77.82) and establishing a solid representation among contemporary deep learning frameworks. This outstanding capability demonstrates that our front-end Adaptive Gated Attention Interaction (AGAI) module successfully models complex atmospheric scattering profiles without inducing severe structural blurring or local contrast attenuation common in conventional feature-stacking pipelines.

Furthermore, to scrutinize the framework's blind aesthetic restoration capability and out-of-distribution generalization, we report the NR-IQA performance on the unconstrained, real-world DHQ dataset. Since the DHQ benchmark contains diverse wild hazy images without available ground-truth reference pairs, it serves as a strict crucible for evaluating realistic color fidelity and structural naturalness. Under this unreferenced setup, GCTC-Net achieves the top-tier score of 49.55, outperforming the prior runner-up GFN (49.30) and the competitive MCN (48.93) by a noticeable margin. This prominent blind evaluation victory directly validates the practical efficacy of our back-end non-linear CCM. Unlike prior networks that rely on hierarchical dynamics with spatial adjacency graph propagation, which frequently suffer from catastrophic white-balance shifts and chromatic aberrations, or modern architectures constrained by elementary function fusion limits like EDN (81.66), our CCM enforces bounded local variance stabilization. Consequently, it effectively rectifies color-casting artifacts, ensuring that the final restored images possess exceptional visual naturalness and zero-shot practical applicability in the wild.

Quantitative evaluation

This paper compares GCTC-Net against 14 state-of-the-art methods, including traditional prior-based approaches (DCP, CAP, Meng, AMEF) and deep learning networks (DehazeNet, AOD-Net, GFN, DCPDN, GCANet, DuRN, GridNet, EDN, FD-GAN, MSBDN). As tabulated in Table 3, GCTC-Net achieves superior performance across all benchmarks. Notably, on the challenging non-homogeneous NH-HAZE dataset, our model scores a PSNR of 20.41 dB and an SSIM of 0.704, outperforming the runner-up (DuRN) by a large margin. On O-HAZE, GCTC-Net likewise secures top-tier results with 25.10 dB PSNR and 0.789 SSIM. Crucially, our model yields the lowest LPIPS scores (0.282 on NH-HAZE and 0.271 on O-HAZE in Table 4).

Table 3. Quantitative comparison on the NH-HAZE dataset.

Method	NH-HAZE		
	PSNR↑	SSIM↑	LPIPS↓
DCP	12.60	0.471	0.465
CAP	13.01	0.445	0.556
Meng	13.12	0.492	0.499
AMEF	13.82	0.506	0.509
DehazeNet	11.76	0.412	0.563
AOD-Net	10.97	0.369	0.654
GFN	12.60	0.388	0.722
DCPDN	11.32	0.467	0.538
GCANet	14.01	0.490	0.491
DuRN	19.38	0.681	0.318
GridNet	17.76	0.644	0.327
EDN	18.58	0.630	0.303
FD-GAN	14.56	0.500	0.480
MSBDN	19.18	0.669	0.307

Method	NH-HAZE		
	PSNR↑	SSIM↑	LPIPS↓
GCTC-Net (Ours)	20.41	0.704	0.282

Table 4. Quantitative comparison on the O-HAZE dataset.

Method	O-HAZE		
	PSNR↑	SSIM↑	LPIPS↓
DCP	14.30	0.587	0.438
CAP	17.46	0.660	0.391
Meng	14.43	0.594	0.495
AMEF	17.30	0.632	0.389
DehazeNet	16.21	0.666	0.386
AOD-Net	10.97	0.369	0.654
GFN	22.58	0.737	0.393
DCPDN	22.78	0.742	0.375
GCANet	24.10	0.758	0.272
DuRN	23.23	0.753	0.315
GridNet	23.67	0.746	0.300
EDN	24.21	0.753	0.274
FD-GAN	16.34	0.574	0.428
MSBDN	23.97	0.760	0.306
GCTC-Net (Ours)	25.10	0.789	0.271

Note: An End-to-End System for Single Image Haze Removal (DehazeNet); All-in-One Dehazing Network (AOD-Net); Gated Fusion Network (GFN); Densely Connected Pyramid Dehazing Network (DCPDN); Gated Context Aggregation Network (GCANet); Dual Residual Network (DuRN); GridDehazeNet: Attention-based Multi-Scale Network for Image Dehazing (GridNet); Ensemble Dehazing Network (EDN); Fusion-discriminator Generative Adversarial Network (FD-GAN); Multi-Scale Boosted Dehazing Network (MSBDN).

This outstanding perceptual advantage demonstrates that our network effectively eliminates white-balance shifts and chromatic distortions, ensuring optimal color fidelity and structural clarity. Fundamentally, while standard convolutional and multi-scale frameworks tend to suffer from localized receptive fields that fail to preserve global structural consistency under non-uniform scattering, GCTC-Net circumvents this limitation by formulating haze removal as a non-local structural modeling task. Inspired by the topological reasoning paradigms of space-time region graphs and the geometric coordinate tracking found in spatial-aware graph relation learning, our bi-dimensional graph network dynamically builds long-range contextual dependencies across spatial and channel dimensions [24,25]. Consequently, it successfully rectifies non-homogeneous degradation patterns and macro-scale chromatic aberrations, providing a mathematical justification for our substantial empirical gains in both low-level fidelity and high-level human perceptual alignment.

Analysis and visualization

Qualitative visual comparisons on the non-homogeneous NH-HAZE dataset are illustrated in Figure 4. Existing CNN-based methods, such as AOD-Net and FFA-Net, fail to resolve uneven haze distributions, leaving dense cloud-like residuals in heavily contaminated regions or causing over-enhancement in thin haze zones. Furthermore, executing extensive deep propagation or deploying unconstrained graph-based global reasoning networks without explicit local boundary constraints systematically induces severe structural blurring and global white-balance shifts [26]. In contrast, GCTC-Net restores sharp background textures and superior color realism, as the AGAI module adaptively filters non-uniform degradations without amplifying background noise.

Figure 4. Qualitative visual comparison of different ablation configurations. Left to right: hazy input, baseline (model 1), baseline + GCP (model 2), baseline + GCP + AGAI (model 3), full model (model 4: baseline + GCP + AGAI + CCM), and ground truth.

The visualization results on the homogeneous O-HAZE dataset in Figures 5 and 6. Although the compared methods successfully remove the overall haze veil, they frequently introduce chromatic aberrations and unnatural color casts, resulting in a faded or oversaturated appearance. Fundamentally, even when utilizing powerful deep residual learning frameworks to stabilize feature propagation and structural reconstruction, these architectures still suffer from cumulative errors in chromaticity under continuous deep layer stacking [27]. To rectify these cumulative errors and balance the feature representations, our back-end non-linear CCM executes localized color variance restoration. Consequently, GCTC-Net suppresses white-balance shifts, producing dehazed images highly consistent with the ground truth in both structural crispness and ambient illumination.

Figure 5. Visual comparison on the NH-HAZE dataset. Left to right: hazy input, DCP, AOD-Net, GCANet, FD-GAN, MSBDN, ours, and ground truth.

Figure 6. Visual comparison of different algorithms on the real-world O-HAZE test dataset. Left to right: hazy input, DCP, AOD-Net, GCANet, FD-GAN, MSBDN, ours, and ground truth.

Quantitative analysis with blind measures

To rigorously evaluate the practical applicability of the proposed GCTC-Net in real-world de-graded scenarios where perfect ground-truth references are fundamentally unavailable. We introduce a dedicated quantitative evaluation based on four widely recognized blind image quality assessment (BIQA) measures: the rate of new visible edges (e), the restoration quality of gradient contrast ($\bar{r}$), the percentage of completely black or saturated pixels (Σ), and the fog awareness density evaluator (FADE). These non-reference metrics provide an objective orthogonal dimension to traditional full-reference criteria, shifting the focus towards perceptual structure visibility, dynamic range fidelity, and absolute haze clearing efficiency.

The blind indicators e and $\bar{r}$ primarily reflect the network's capacity to reconstruct lost high-frequency spatial details and local contrast profiles. A higher value of e signifies that GCTC-Net successfully unveils previously obscured structural boundaries and fine textures under non-homogeneous scattering conditions, while an elevated $\bar{r}$ indicates that the sharpened gradients look natural and free of excessive ringing artifacts. Concurrently, the metric Σ is leveraged to monitor the percentage of over-enhanced, under-exposed, or severely clipped pixels. Minimizing Σ is paramount for non-homogeneous dehazing, as aggressive transmission estimation often leads to unnatural dark shadows or blown-out highlights. By strictly containing Σ, we empirically demonstrate that our bounded residual CCM and adaptive gating constraints effectively regularize the dynamic range, preventing structural over-saturation.

A detailed quantitative diagnostic on the City Bund dataset reveals insightful behaviors across different paradigms. Traditional prior-based methods such as Meng and DCP exhibit remarkably high scores in edge product indicators (e and $\bar{r}$), which is mathematically attributed to their heavy reliance on local min-filtering and extreme patch-wise contrast stretching. However, this aggressive structural sharpening often comes at the cost of potential noise amplification in complex real-world scenes. In sharp contrast, early learning-based frameworks (e.g., AOD-Net and DehazeNet) struggle to uncover blocked visibility under severe non-homogeneous degradation, yielding subdued e values below 2.000. Meanwhile, generalized multi-scale models like MSBDN deliver a suboptimal FADE score of 3.749, indicating prominent haze residues.

Distinct from these extremes, our GCTC-Net yields a highly balanced and competitive performance profile. Specifically, GCTC-Net secures a solid visible edge rate of 4.415 and a gradient contrast score of 2.342, significantly outperforming contemporary powerful baseline architectures such as GFN and MCN. More importantly, while successfully preserving high-frequency structural boundaries, our model suppresses structural clipping to absolute zero (Σ=0) and fundamentally drives the residual haze density down to an elite level (FADE: 1.258) in Table 5. As summarized via these indicators, this multi-faceted BIQA superiority firmly demonstrates that GCTC-Net does not merely over-fit to pixel-wise ground-truth statistics, but intrinsically restores highly clear, perceptually natural, and structural-crisp scene representations from severe non-homogeneous atmospheric degradation.

Table 5. Quantitative comparison of blind image quality assessment (BIQA) measures on the City Bund dataset. The best and second-best results are highlighted in bold and underlined, respectively.

Method	City Bund			
	$e\uparrow$	$\Sigma\downarrow$	$\bar{r}\uparrow$	FADE $\downarrow$
DCP	9.899	0	4.110	0.428
CAP	2.353	0	1.363	1.397
Meng	11.556	0	4.626	0.404
AMEF	3.291	0	2.255	1.424
DehazeNet	1.852	0	1.525	2.059
AOD-Net	1.435	0	1.497	3.189
GFN	4.276	0	2.348	0.873
DCPDN	2.496	0	1.954	2.123
GCANet	3.685	0	2.195	1.353
DuRN	4.008	0	2.093	1.765
GridNet	3.347	0	1.939	2.144
EDN	2.940	0	1.806	2.231
FD-GAN	7.441	0	3.420	0.537
MSBDN	0.855	0	1.407	3.749
MCN	4.258	0	2.327	1.361
GCTC-Net	4.415	0	2.342	1.258

Ablation study

To verify each component's contribution in GCTC-Net, ablation studies are performed on the NH-HAZE dataset under four distinct configurations (Table 6), with corresponding qualitative evolution sequences visualized in Figure 4. This choice of evaluation vehicle is motivated by the fact that while conventional synthetic fog removal often relies on simple domain or model adaptation strategies, resolving highly non-homogeneous atmospheric scattering - as formulated in the landmark new trends in image restoration and enhancement (NTIRE) benchmarks - demands sophisticated spatial and prior reasoning [28].

Model 1 establishes the pure backbone baseline without illumination guidance. As observed in Figure 4, because the backbone relies entirely on uniform receptive fields to estimate deep mappings, it exhibits limited adaptation to non-uniform haze density variations, resulting in substantial smoke residues and local contrast attenuation. In model 2, inspired by con-temporary enhancement frameworks that fuse rich intrinsic prior information to regularize ill-posed restoration, the generated contrastive pseudo prior is introduced via a rigid concatenation strategy (torch.cat followed by a 1×1 convolution), yielding a marginal PSNR improvement from 17.15 dB to 17.85 dB [29]. Visually, while model 2 reconstructs a primitive scene profile and partially recalibrates the coarse transmission, the rigid feature stacking injects considerable prior redundancy into the deep layers, which in turn leads to severe over-saturation in thin-haze regions and artifacts around thick haze boundaries.

Crucially, by integrating the Adaptive Gated Attention Interaction (AGAI) module (GF_layer) in Model 3, the PSNR significantly jumps to 19.52 dB. This empirical and visual boost demonstrates that our gated mechanism successfully filters feature redundancy and adaptively distills beneficial illumination-enhanced priors over blind feature stacking. As qualitatively evidenced in the fourth column of Figure 4, model 3 eliminates the dense cloud-like smoke pockets without incurring background noise amplification or halo artifacts, re-covering exceptionally sharp structural details and background contours. Evaluating the non-linear CCM by comparing model 3 with model 4 (full model) reveals a vital insight. Although model 3 achieves competitive structural restoration, its perceptual color fidelity remains suboptimal (LPIPS: 0.332).

This is because long-range dependencies in deep graph propagation tend to over-propagate global illumination components, which inadvertently amplifies cross-channel deviations and induces global white-balance shifts or unnatural color casts (e.g., subtle yellow-green tinting across neutral surfaces). By activating the back-end CCM, model 4 achieves the optimal LPIPS score of 0.282 while further fine-tuning PSNR and SSIM. This striking perceptual improvement is visually striking in the fifth column of Figure 4: The bounded residual color calibration network provides a robust locally stabilized constraint to suppress chromatic aberrations and restore authentic ambient illumination. The final output of the Full model yields an elite level of color realism and structural crispness, aligning almost perfectly with the ground truth.

Table 6. Quantitative ablation results of different module combinations on the NH-HAZE dataset.

Model	Baseline	GCP Prior	GF_layer (AGAI)	CCM	PSNR (dB)↑	SSIM↑	LPIPS↓
1	✓	/	/	/	17.15	0.621	0.388
2	✓	✓	/	/	17.85	0.640	0.365
3	✓	✓	✓	/	19.52	0.689	0.332
4 (Full)	✓	✓	✓	✓	20.41	0.704	0.282

Discussion

Through extensive quantitative and qualitative evaluations, we have thoroughly validated the efficacy of incorporating generative contrastive pseudo priors and dual-path topo-logical graph reasoning for non-homogeneous haze removal. The synergy between the front-end AGAI module and the back-end CCM establishes a highly robust pipeline capable of adaptively modeling non-uniform atmospheric scattering and stabilizing local color variances across diverse, open-world scenes.

Nevertheless, while our GCTC-Net achieves state-of-the-art restoration performance on static unconstrained images, a critical deployment challenge emerges when extending this framework to continuous video dehazing streams. Specifically, the framework executes deep graph reasoning and non-linear color calibration independently for each individual frame. When handling high-frame-rate video sequences, the profound spatiotemporal correlation and structural redundancy between adjacent frames remain unexploited, which systematically elevates the overall computational overhead and may induce subtle inter-frame flickering artifacts due to the independent local variance stabilization in CCM.

Conclusion

In this paper, GCTC-Net is introduced as a novel graph convolutional texture and color calibration network optimized for non-homogeneous image dehazing. By utilizing local graph convolutions alongside the generated contrastive pseudo prior, the model successfully captures long-range spatial dependencies and introduces virtual multi-exposure channels to overcome the limitations of single-image inputs. Furthermore, the Adaptive Gated Attention Interaction (AGAI) module (GF_layer) effectively distills illumination-guided representations while suppressing feature redundancy, and the non-linear CCM applies a robust bound residual constraint to rectify cross-channel white-balance drifts. Extensive evaluations on the NH-HAZE dataset demonstrate that our method achieves superior dehazing and color-rectifying performance. In our future work, we will focus on light-weighting the architecture via network pruning to enhance its deployment efficiency on resource-constrained edge devices, and explore its generalization capabilities in other degraded visionary scenarios such as underwater image restoration.

Funding

This work was not supported by any funds.

Acknowledgements

The authors would like to show sincere thanks to those techniques who have contributed to this research.

Conflicts of Interest

The authors declare no conflict of interest.

References

[1] Yuan, H., Yan, T. H. (2024) Parallel multi scale attention mapping image dehazing algorithm. *Laser & Optoelectronics Progress*, 61(4), 101-112.

[2] Wang, Q., Jiang, K., Wang, Z., Ren, W., Zhang, J., Lin, C. W. (2023) Multi-scale fusion and decomposition network for single image deraining. *IEEE Transactions on Image Processing*, 33, 191-204.

[3] Minaee, S., Boykov, Y., Porikli, F., Plaza, A., Kehtarnavaz, N., Terzopoulos, D. (2021) Image segmentation using deep learning: a survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(7), 3523-3542.

[4] Zhou, R., Wang, R., Wang, Y., Zhou, F., Guo, N. (2024) Research on low-light image enhancement based on MER-Retinex algorithm. *Signal, Image and Video Processing*, 18(1), 803-811.

[5] Othman, G., Zeebaree, D. Q. (2020) The applications of discrete wavelet transform in image processing: a review. *Journal of Soft Computing and Data Mining*, 1(2), 31-43.

[6] Lisani, J. L., Morel, J. M., Petro, A. B., Sbert, C. (2020) Analyzing center/surround retinex. *Information Sciences*, 512, 741-759.

[7] Singh, D., Kumar, V. (2018) Single image haze removal using integrated dark and bright channel prior. *Modern Physics Letters B*, 32(04), 1850051.

[8] Zhu, Q., Mai, J., Shao, L. (2015) A fast single image haze removal algorithm using color attenuation prior. *IEEE Transactions on Image Processing*, 24(11), 3522-3533.

[9] Huang, Z., Jing, H., Chen, A., Hong, C., Shang, X. (2023) Efficient image dehazing algorithm using multiple priors constraints. *Journal of Visual Communication and Image Representation*, 90, 103694.

[10] Yang, A., Liu, Y., Wang, J., Li, X., Cao, J., Ji, Z., Pang, Y. (2023) Visual-quality-driven unsupervised image dehazing. *Neural Networks*, 167, 1-9.

[11] Zhu, X., Ning, J., Wu, X., Duo, B. (2025) Frequency attention-based multi-scale dehazing GAN. *Digital Signal Processing*, 105747.

[12] Li, J., Zheng, M., Sun, C., Wang, Z., Shen, H. (2026) T-SFANet: A spatial-frequency attention network for thermal infrared image dehazing. *Neurocomputing*, 132797.

[13] Cai, B., Xu, X., Jia, K., Qing, C., Tao, D. (2016) Dehazenet: An end-to-end system for single image haze removal. *IEEE Transactions on Image Processing*, 25(11), 5187-5198.

[14] Chen, Z., He, Z., Lu, Z. M. (2024) DEA-Net: Single image dehazing based on detail-enhanced convolution and content-guided attention. *IEEE Transactions on Image Processing*, 33, 1002-1015.

[15] Batool, I., Imran, M. (2025) A dual residual dense network for image denoising. *Engineering Applications of Artificial Intelligence*, 147, 110275.

[16] Hu, B. (2021) Multi-scale feature fusion network with attention for single image dehazing. *Pattern Recognition and Image Analysis*, 31(4), 608-615.

[17] Park, J., Han, D. K., Ko, H. (2020) Fusion of heterogeneous adversarial networks for single image dehazing. *IEEE Transactions on Image Processing*, 29, 4721-4732.

[18] Zheng, L., Li, Y., Zhang, K., Luo, W. (2022) T-Net: Deep stacked scale-iteration network for image dehazing. *IEEE Transactions on Multimedia*, 25, 6794-6807.

[19] Zhang, X., Song, M., Fan, Z., Jin, H. (2022) Reconstruction algorithm of haze image based on blind separation model of polarized orthogonal airlight. *Optics Express*, 30(23), 42097-42113.

[20] Hou, Y., Xu, J., Liu, M., Liu, G., Liu, L., Zhu, F., Shao, L. (2020) NLH: A blind pixel-level non-local method for real-world image denoising. *IEEE*

Transactions on Image Processing, 29, 5121-5135.

[21] Dong, P., Wang, B. (2022) TransRA: Transformer and residual attention fusion for single remote sensing image dehazing. *Multidimensional Systems and Signal Processing*, 33(4), 1119-1138.

[22] Li, B., Ren, W., Fu, D., Tao, D., Feng, D., Zeng, W., Wang, Z. (2018) Benchmarking single-image dehazing and beyond. *IEEE Transactions on Image Processing*, 28(1), 492-505.

[23] Sakaridis, C., Dai, D., Van Gool, L. (2018) Semantic foggy scene understanding with synthetic data. *International Journal of Computer Vision*, 126(9), 973-992.

[24] Nicolicioiu, A., Duta, I., Leordeanu, M. (2019) Recurrent space-time graph neural networks. *Advances in Neural Information Processing Systems*, 32.

[25] Chen, S., Li, Z., Tang, Z. (2020) Relation R-CNN: A graph based relation-aware network for object detection. *IEEE Signal Processing Letters*, 27, 1680-1684.

[26] Wang, R., Huang, C., Wang, X. (2020) Global relation reasoning Graph Convolutional Networks for human pose estimation. *IEEE Access*, 8, 38472-38480.

[27] Shafiq, M., Gu, Z. (2022) Deep residual learning for image recognition: a survey. *Applied Sciences*, 12(18), 8972.

[28] Dai, D., Sakaridis, C., Hecker, S., Van Gool, L. (2020) Curriculum model adaptation with synthetic and real data for semantic foggy scene understanding. *International Journal of Computer Vision*, 128(5), 1182-1204.

[29] Li, Y., Wei, X., Liao, X., Zhao, Y., Jia, F., Zhuang, X., Zhou, M. (2024) A deep retinex-based low-light enhancement network fusing rich intrinsic prior information. *ACM Transactions on Multimedia Computing, Communications and Applications*, 20(11), 1-23.