

ORIGINAL RESEARCH OPEN ACCESS

Adaptive Convolutional Strategy for Robust Image Dehazing in Diverse Environments

Hira Khan¹ | Sung Won Kim² 
¹Department of Information and Communication Engineering, Yeungnam University, Gyeongsan, Republic of Korea | ²School of Computer Science and Engineering, Yeungnam University, Gyeongsan, Republic of Korea

Correspondence: Sung Won Kim (swon@yu.ac.kr)

Received: 17 March 2025 | **Revised:** 21 August 2025 | **Accepted:** 8 September 2025

Funding: This research was supported in part by the Basic Science Research Program through the National Research Foundation of Korea (NRF), funded by the Ministry of Education (NRF-2021R1A6A1A03039493) and in part by the NRF grant funded by the Korean government (MSIT) (NRF-2022R1A2C1004401).

ABSTRACT

Adverse weather conditions such as haze, fog, and smog degrade image visibility, adversely affecting the performance of vision-based systems. Existing dehazing methods often struggle with non-uniform haze distributions, limited detail restoration, and poor generalization across diverse scenes. To overcome these limitations, this paper presents a deep learning-based dehazing framework that jointly restores image clarity and detail. Unlike conventional algorithms that often neglect fine structure recovery, our architecture incorporates four specialized sub-modules: (i) a noise attention module for enhancing noise suppression and feature preservation; (ii) an adaptive ConvNet module; (iii) a feature extraction module for capturing salient image features; and (iv) a detail refinement module to enhance spatial fidelity. The architecture is trained in an end-to-end manner to restore both structural integrity and colour consistency under challenging conditions. Extensive experiments conducted on synthetic and real-world datasets, including indoor, outdoor, underwater, night-time, and remote sensing scenarios, demonstrate superior generalization capability. In the SOTS indoor dataset, our method achieves a PSNR of 28.44 dB and an SSIM of 0.967, outperforming several state-of-the-art methods. Evaluations using additional metrics such as CIEDE2000 and MSE confirm the effectiveness of the proposed method in handling dense and heterogeneous haze while preserving fine textures and visual fidelity.

1 | Introduction

Haze and related atmospheric conditions degrade image quality by reducing contrast and visibility, significantly affecting both low- and high-level computer vision tasks [1]. To address this, a variety of dehazing techniques have been proposed, including contrast enhancement [2], prior-based [3], and, more recently, deep learning-based [4]. Among these, data-driven single image dehazing (SID) methods have shown great promise due to their ability to learn complex mappings directly from data without relying on hand-crafted priors. However, existing methods often struggle in non-uniform haze conditions and fail to

preserve fine details, highlighting the need for more robust and generalizable solutions.

The structure of the physical scattering model is shown in Figure 1. It can be expressed as follows.

$$F(x) = G(x) \cdot T(x) + \gamma \cdot (1 - T(x)) \quad (1)$$

where $G(x)$ represents the clear image, $T(x)$ indicates the transmission coefficient, and γ signifies the atmospheric factor. Upon examining the equation above, it is evident that the only familiar element is the hazy image $F(x)$, while the other elements

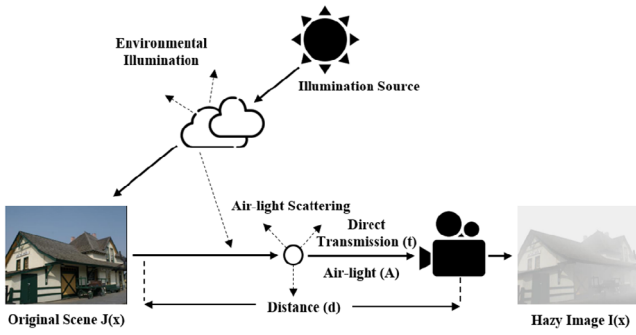


FIGURE 1 | Atmospheric scattering model.

represent unknown terms. The computation of the transmission factor is expressed as

$$T = e^{-\delta d} \quad (2)$$

where δ signifies the scattering coefficient and d denotes the range between the scene and the sensors. Subsequently, the restored image can be described as

$$G(x) = \frac{F(x) - B}{T(x)} \quad (3)$$

Estimating the transmission map $T(x)$ and atmospheric light B is central to image dehazing but remains an ill-posed challenge due to their inherent uncertainty. To overcome this, prior-based single image dehazing (SID) methods impose constraints derived from image statistics or physical assumptions. Two major traditional approaches have emerged: contrast enhancement and prior-based techniques. Contrast-based method [2] improve visibility by boosting brightness and contrast but risk information loss in brightened regions.

In addition, prior-based techniques have been proposed that exploit statistical or physical assumptions, such as dark channel, wavelet transform, and colour attenuation priors [5]. These methods often outperform contrast-based techniques, but typically involve higher computational complexity.

Recent advances in deep learning have led to significant improvements in visual perception tasks [6]. Learning-based SID techniques have emerged [7], leveraging convolutional neural networks (CNNs) to predict transmission maps or directly restore clean images.

DehazeNet [8] estimates transmission using a CNN-based architecture, while MSCNN [9] performs multi-scale transmission prediction. Holistic networks such as GridDehazeNet [10] bypass intermediate estimations by learning a direct mapping from hazy to clear images, offering improved restoration.

Although these methods differ in structure, they share the fundamental goal of estimating the unknowns in the atmospheric scattering model (e.g. B and $T(x)$) to recover the clean image $G(x)$ as described in Equation (1).

From an estimation theory stand-point, most dehazing techniques follow a module-based design, attempting to restore

clear images by estimating intermediate variables such as B and $T(x)$. However, this strategy has inherent limitations, as the estimation of these parameters does not always guarantee accurate reconstruction of $G(x)$, especially when evaluated using the same objective metrics.

Consequently, plug-in models trained on synthetic datasets often fail to generalize well to real-world conditions, limiting their practical applicability.

Despite significant progress in deep learning-based dehazing, many existing methods operate under the assumption that the haze is uniformly distributed. However, as shown in Figure 2, real-world haze often follows a non-uniform distribution, deviating from the idealized homogeneous model. For instance, the green halo in non-homogeneous haze appears more diffused around the object, contrasting with the sharper boundaries in homogeneous haze scenarios. In homogeneous haze removal, traditional methods often produce dehazed outputs with reduced contrast and loss of fine details. This limitation stems from their inability to establish a robust mapping between hazy and clear images, primarily due to the restricted learning capacity of their core components when faced with complex or non-uniform haze conditions [4]. To provide a holistic view of recent advancements in image dehazing, Table 1 compares state-of-the-art methods and the proposed approach in terms of architecture, preprocessing needs, dataset variety, performance metrics, strengths and limitations. This comparative analysis highlights the domain adaptability, structural innovation, and evaluation rigor of our model. The remaining sections of the paper are structured as follows. Section 2 provides an overview of recent advances in dehazing methods, with a focus on those based on deep learning. Section 3 introduces our proposed model and outlines its key modules. Section 4 elaborates on the datasets, and Section 5 wraps up the findings and explores potential avenues for future research.

1.1 | Motivation and Key Contributions

Haze removal is often accompanied by the loss of important background details, leading to degradation in visual quality. Many existing single-image dehazing methods emphasize haze suppression while overlooking texture refinement, resulting in incomplete restoration. To overcome the limitations of ASM-dependent models and their poor adaptability in real-world heterogeneous haze, this work proposes a flexible, learning-based framework that eliminates the need for handcrafted priors. To address this, the proposed method integrates both dehazing and detail enhancement in a unified framework. It introduces four specialized components: a noise attenuation module (NAM), an adaptive ConvNet module (ACNM), a feature extraction module (FEM), and a detail refinement module (DRM). Each module is designed to handle a distinct restoration task, enabling more effective recovery of both structural clarity and fine textures across diverse scenes.

The proposed model is well suited for real-world applications such as autonomous driving, surveillance, UAV-based imaging, underwater exploration, and satellite observation. Its modular architecture generalizes well across diverse domains, including

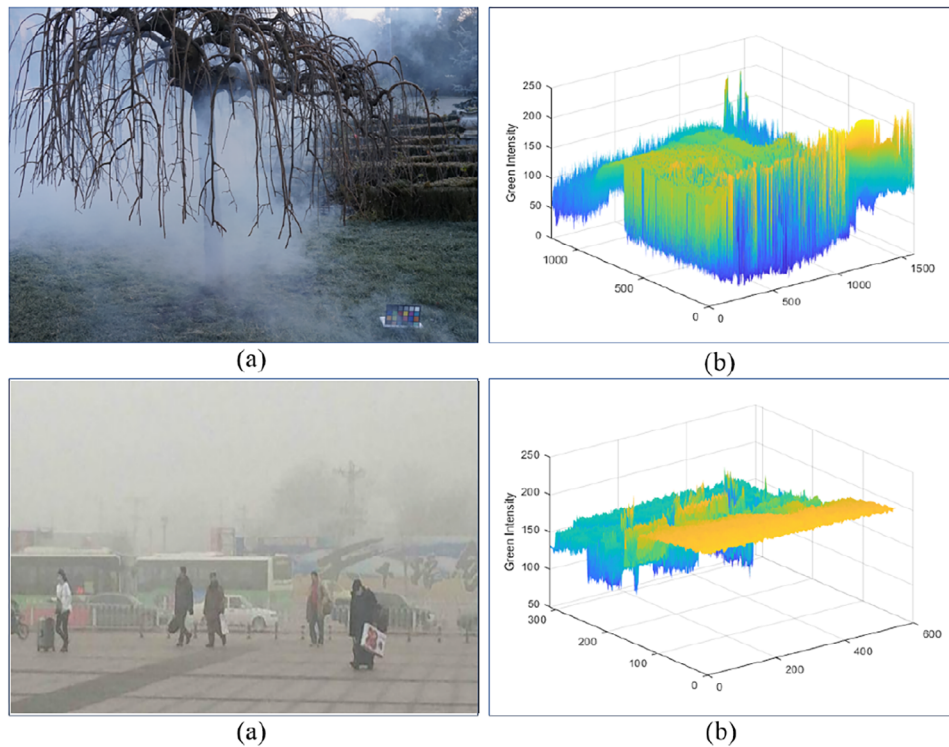


FIGURE 2 | Row 1: (a) non-homogeneous haze (b) 3D green intensity graph of non-homogeneous haze, Row 2: (a) homogeneous haze, (b) 3D green intensity graph of homogeneous haze

daytime, night-time, remote sensing, and underwater scenes, without requiring architecture-specific tuning. Although the current implementation relies on paired data, the model can be extended for semi-supervised training. Its lightweight design also supports the deployment on edge devices with moderate computational resources.

The key contributions of the proposed work are outlined as follows:

1. We propose a novel end-to-end deep learning network comprising four dedicated submodules: NAM, ACNM, FEM, and DRM, each enhancing image clarity and contrast.
2. The ACNM module adopts an encoder-decoder structure that emphasizes informative channel features, thereby boosting dehazing performance.
3. Our model restores haze-free images without estimating transmission maps or atmospheric light, leveraging FEM for multiscale contextual processing.
4. Extensive experiments were conducted on diverse datasets including indoor, outdoor, remote sensing, underwater, and night-time scenes to demonstrate the superiority of our approach over existing methods.

2 | Related Work

Single image dehazing (SID) has been approached through both prior-driven and learning-based methodologies. This section categorizes related work into three groups: hand-crafted

prior-based models, deep learning-based strategies, and recent domain-adaptive or real-time dehazing efforts.

2.1 | Prior-Based Strategies

Traditional SID techniques rely on statistical or physical assumptions about the haze formation process. The dark channel prior (DCP) by He et al. [23] is a foundational method that assumes low-intensity pixels in haze-free images, allowing transmission map estimation. Follow-up works enhanced its robustness [5, 24].

Alternative priors have emerged to estimate scene depth and atmospheric light. Zhu et al. [3] proposed a colour attenuation model based on the brightness and saturation of the pixels, while Fattal [25] used colour line distributions in small patches. Berman et al. [26] introduced non-local clusters of colour lines, and improved pixel separation using colour ellipsoids.

Although effective in some cases, these models are sensitive to their assumptions and may cause artefacts such as colour distortion or inaccurate transmission estimates, especially in complex or dense haze scenes.

2.2 | Learning-Based Strategies

Data-driven SID methods leverage CNNs to overcome the rigidity of hand-crafted priors. DehazeNet [8] estimates transmission using learned filters, while AOD-Net [11] combines transmission and atmospheric light into a unified prediction task.

TABLE 1 | Unified comparative analysis of recent deep learning-based image dehazing models, including our proposed model. All models marked Yes under the Preproc. column apply custom pre-processing such as haze simulation, intensity normalization, or contrast adjustment. Our model uses only basic resizing and normalization to tensor format.

S.	Method	Year	Key characteristics	Preproc.	Dataset	Performance Metrics	Advantages	Limitations
1	DehazeNet [8]	2016	Maxout unit + custom activation	Yes	Middlebury, Synthetic	MSE, SSIM, PSNR, WSNR	Learns transmission map	Doesn't optimize dehazing
2	AOD-Net [11]	2017	Unified ASM model (no separate maps)	No	NYU2, Pascal VOC	PSNR, SSIM, MSE	End-to-end flow	Artefacts in dense haze
3	DehazeFormer [12]	2022	Transformer with normalization and activation tuning	No	RESIDE, RS-Haze, SOTS	PSNR, SSIM	High accuracy ViT	High complexity
4	MFINEA [13]	2023	Multi-feature interaction edge-aware module	Yes	RESIDE, NTIRE, NH-Haze	PSNR, SSIM	Strong edge detail	Weak CNN hierarchy
5	Comp.-neighbourhood [14]	2023	Attenuation coefficient for neighbourhood intensity	Yes	Natural, OTS, ITS	PSNR, SSIM, Haze density, EV	Preserves local contrast	High complexity
6	Dual stream [15]	2023	Frequency + content dual-stream network	No	SOTS, I-Haze, O-Haze, Dense-Haze	PSNR, SSIM	Preserves edge texture	Weak on SOTS indoor
7	DenseConvTrans [16]	2023	Convolutional transformer with metric learning	No	RESIDE, NH-Haze, Dense-Haze	PSNR, SSIM	Joint loss fusion	Overfits synthetic data
8	LapDehazeNet [17]	2023	Taylor-based Laplacian pyramid	Yes	4KID, O-Haze, I-Haze	PSNR, SSIM, FLOPS	UHD enhancement	High runtime, not mobile-suitable
9	Deep hybrid [4]	2023	Treats dehazing and refinement as separate modules	No	RESIDE, NH-Haze, SOTS	PSNR, SSIM, LPIPS, DHQI, FRFSIM	Refined results	Not for real-time use
10	Wavelet AE [18]	2024	Wavelet/inverse wavelet transform in AE	No	RESIDE-6K, Rain800, ITS, OTS	PSNR, SSIM	Frequency-aware up/down sampling	Requires GT supervision
11	DEA-Net [19]	2024	DEConv + CGA guided attention blocks	Yes	RESIDE (OTS/ITS), RTTS, HSTS	PSNR, SSIM	Better feature maps	Needs paired aligned data
12	UCL-Dehaze [20]	2024	Contrastive + adversarial training	No	RESIDE, SOTS, RTTS, HSTS, URHI	PSNR, SSIM, CIEDE2000	Works without GT	Instability, GAN tuning
13	DR3DF-Net [21]	2025	3D feature routing with dynamic constraints	No	RSID, RICE, SateHazelk	PSNR	Robust cross-domain routing	Weak for thick haze
14	DNMGDT [22]	2025	Multi-prior + adaptive loss domain transfer	No	SOTS, I-Haze, O-Haze, MARFID	PSNR, SSIM, IQI, FRFSIM	Adapts to multiple domains	Struggles with colour
15	Proposed model	2025	Hybrid CNN-transformer with NAM, ACNM, FEM, DRM for detail-aware dehazing	No	Indoor, Outdoor, Remote, Night, Underwater	PSNR, SSIM, CIEDE2000, MSE	Joint dehazing and detail enhancement	Requires paired data

Ren et al. [27] introduced the gated fusion network (GFN), which combines various enhanced images to generate haze-free outputs without relying on atmospheric priors. Shao et al. [28] focused on domain adaptation to bridge synthetic-real domain gaps.

Recent architectures adopt encoder–decoder structures for direct image restoration. MFINEA [13] fuses hierarchical features to enhance edges and [4] uses residual attention modules to recover fine details. Pyramid CNNs and U-Net variants [29] also contribute to the preservation of spatial fidelity.

Another direction leverages weak supervision: Fahim et al. [30] proposed a semisupervised atmospheric component learning method using no-reference metrics and physical priors to restore visibility under low-light conditions. These approaches reflect the trend towards attention-based and self-supervised strategies, which inspire the design and training choices in our work.

2.3 | Modular and Comparative Strategies

Some models, such as LIDN [31] employ modular designs similar to ours, yet struggle with issues such as colour shifts. Others, like GFN and [32] eliminate reliance on the atmospheric scattering model (ASM), improving generalization. However, these models either lack structural flexibility or require pre-processing schemes.

In contrast, our model introduces a fully modular, ASM-independent hybrid framework comprising noise suppression, adaptive encoding, feature aggregation, and detail refinement. This enhances robustness across varied haze densities and domains without additional pre-processing or hand-crafted priors.

2.4 | Application-Oriented Strategies

Recent studies have introduced additional innovations in deep learning-based dehazing. For example, Wang et al. [33] employed adaptive Retinex and prior fusion to improve visibility under extreme conditions. Alenezi et al. [34] proposed a real-time dehazing algorithm optimized for low-latency environments.

Lian et al. [35] presented a feature fusion approach with multi-scale attention to better handle noise and varying illumination. Hu et al. [36] isolates the sky region of hazy images using mean shift with sky color priors through adaptive thresholding based on that region. Sahu et al. [37] introduced a multistream design tailored for intelligent transportation. Compared to these task-specific frameworks, our method is designed to generalize across diverse real-world domains without requiring application-specific tuning.

3 | Proposed Method

The proposed framework comprises four submodules: (i) noise attenuation module (NAM) for noise suppression; (ii) adaptive ConvNet module (ACNM) for encoder–decoder based dehazing; (iii) detail refinement module (DRM) for structural and textural

enhancement; and (iv) feature extraction module (FEM) for local feature recovery. Each module performs a distinct sub-task, and their outputs are fused via a learnable weighted summation to produce the final dehazed result.

Figure 3 illustrates the modular architecture, where all submodules receive the input image in parallel. This design promotes complementary feature learning without sequential redundancy.

3.1 | Noise Attenuation Module

The NAM reduces sensor and atmospheric noise while preserving essential details. It applies successive convolutions with tanh activation to maintain gradient smoothness and batch normalization (BN) for stable convergence. The early layers X_1 to X_3 (with 128 filters) capture noise patterns, while the deeper layers X_6 to X_9 refine the structure, producing the denoised output O_{NAM} .

Let's $I \in \mathbb{R}^{H \times W \times C}$ denote the input image with $H = 400$, $W = 400$, and $C = 3$. The convolutional operation is defined as:

$$X_i = f(W_i * X_{i-1} + b), \quad i = 1, 2, 3, 4 \quad (4)$$

where W_i and b_i represent the convolution filters and biases, and $f(\cdot)$ is the tanh activation function:

$$f(x) = \tanh(x) \quad (5)$$

Batch normalization is applied as

$$N_1 = \frac{X_4 - \mu_{X_4}}{\sigma_{X_4} + \epsilon} \gamma + \beta \quad (6)$$

where μ_{X_4} and σ_{X_4} are the mean and standard deviation of X_4 , and γ, β are learnable parameters.

After deeper convolutions (X_6 to X_9), the final output of the NAM is

$$O_{\text{NAM}} = N_2 \quad (7)$$

3.2 | Adaptive ConvNet Module

The ACNM Module serves as a generative encoder–decoder network for end-to-end image dehazing. The term “adaptive” refers to the module’s ability to dynamically adjust its feature processing pipeline to varying haze distributions. Skip connections allow the reuse of both low-level spatial features and high-level abstractions, enabling the network to generalize across diverse haze conditions without relying on handcrafted priors.

This design preserves structural fidelity and fine textures while effectively removing haze. The FEM is applied after the detail refinement module, further encodes residual spatial features to enhance structure, texture, and localized details in the final reconstruction stage.

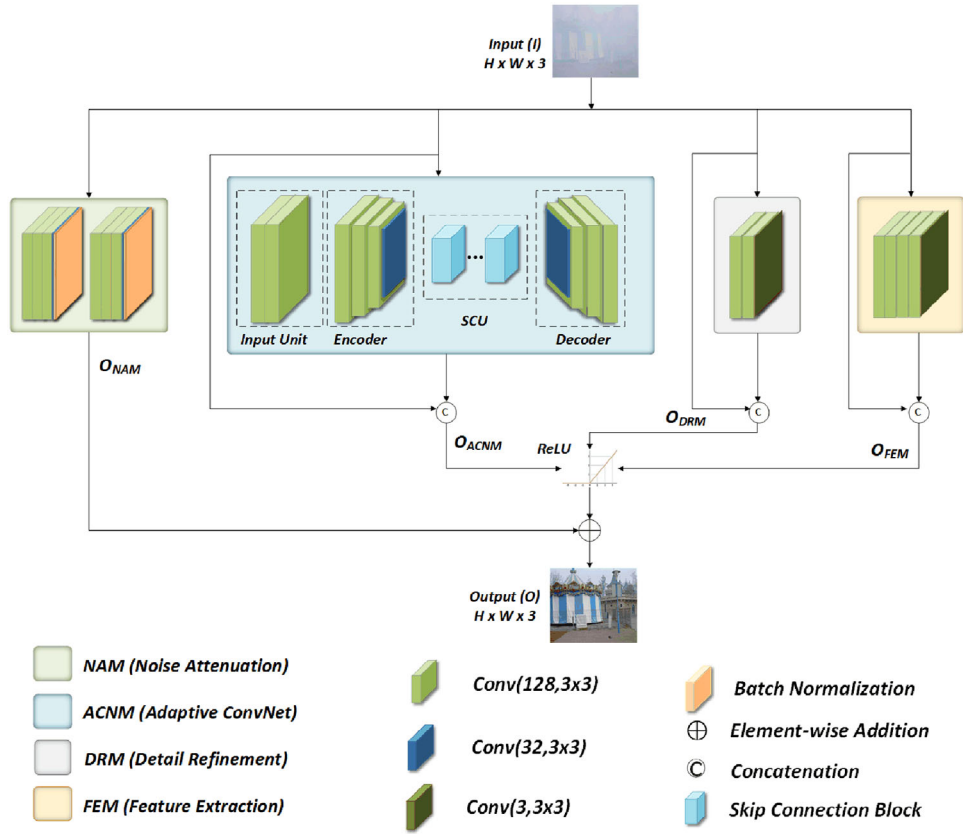


FIGURE 3 | Overall architecture of the proposed modular dehazing framework. The four modules: noise attenuation, adaptive ConvNet, detail refinement, and feature extraction-process the input image in parallel, and their outputs are fused through a learnable weighted summation to generate the final dehazed image. Colour coding indicates functional categories as shown in the legend.

3.2.1 | Encoding Phase

The input image I undergoes a series of convolutional operations that reduce spatial resolution and capture abstract features. The transformation begins with:

$$U_{X_2} = f(W_0 * I + B_0) \quad (8)$$

Here, $f(\cdot)$ denotes a non-linear activation function, which enables the capture of complex image patterns and high-frequency details. Successive encoder layers are defined as:

$$E_{X_i} = f(W_{E_i} * E_{X_{i-1}} + b_{E_i}), \quad i = 1, 2, 3 \quad (9)$$

Here, E_i denotes the i th encoder feature map generated during the encoding stage of the ACNM Module, which progressively captures abstract image representations at multiple levels. The final encoder layer, with 32 filters, compresses the feature maps to a compact representation:

$$E_{z_1} = f(W_{z_1} * E_{X_3} + b_{z_1}) \quad (10)$$

3.2.2 | Skip Connection Blocks (SCBs)

To preserve spatial information and improve gradient flow, skip connection blocks (SCBs) are incorporated. These

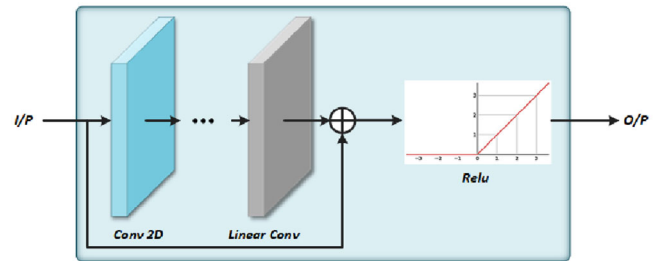


FIGURE 4 | The skip connection block architecture. The number of convolutional layers could be different for three different blocks (3, 3, and 5). The detailed description of each block is explained in Section 3.

blocks directly pass encoder features to the corresponding decoder layers, helping retain fine-grained details. The detailed structure of the skip connection block is illustrated in Figure 4.

Each SCB consists of convolutional operations and ReLU activation:

$$SCB_i = g(W_{SCB_i} * E_{X_i} + b_{SCB_i}) + E_{X_i} \quad (11)$$

where $g(x) = \max(0, x)$ is the ReLU activation function.

3.2.3 | Decoding Phase

The decoder reconstructs the dehazed image through transposed convolutions, gradually increasing spatial resolution. Initial layers restore global structure, while deeper layers refine textures.

$$D_{z_i} = f(W_{D_i} * SCB + b_{D_i}), \quad i = 1, 2, 3, 4 \quad (12)$$

The final dehazed output is obtained via:

$$O_{ACNM} = \sigma(W_{\text{final}} * D_{X_2} + b_{\text{final}}) \quad (13)$$

Here, it $\sigma(x) = \text{sigmoid}(x)$ ensures pixel values remain within the valid range [0,1].

3.3 | Detail Refinement Module

The DRM enhances edges and fine textures. It applies a 128-filter ReLU convolution for salient structures, followed by tanh layers for smooth texture recovery. To preserve natural appearance, the refined features are concatenated with the original input:

$$D_{X_1} = g(W_{D_1} * I + b_{D_1}) \quad (14)$$

In this context, D_i represents the i th decoder feature map within the DRM module, which emphasizes significant patterns and progressively enhances fine structural details. Subsequent layers use $\tanh$ activation to smooth transitions and capture fine-grained texture.

$$D_{X_2} = \tanh(W_{D_2} * D_{X_1} + b_{D_2}) \quad (15)$$

$$D_{X_3} = W_{D_3} * D_{X_2} + b_{D_3} \quad (16)$$

The final output combines the refined features with the original input to avoid over-modification:

$$D_{\text{DRM}} = \text{Concat}(I, D_{X_3}) \quad (17)$$

3.4 | Feature Extraction Module

The FEM Module enhances local contrast and edge details by learning hierarchical representations. It is applied after the DRM to capture residual structural and textural information, improving the final reconstruction.

To achieve this, it uses three convolutional layers with ReLU activation, followed by a final linear convolution to condense features:

$$F_{X_i} = g(W_{F_i} * F_{X_{i-1}} + b_{F_i}), \quad i = 1, 2, 3 \quad (18)$$

Here, F_i corresponds to the i th hierarchical feature map extracted in the FEM module, responsible for learning residual spatial and textural information critical for detail preservation. A final convolutional layer transforms the output into a condensed feature representation:

$$F_{X_4} = W_{F_4} * F_{X_3} + b_4 \quad (19)$$

ALGORITHM 1 | Proposed dehazing method.

Input: Hazy image I , initialized parameters θ

Output: Enhanced dehazed image O_{final}

- 1 Initialize θ for NAM, ACNM, FEM, DRM, learning rate η , iteration counter $t = 0$;
- 2 Extract noise patterns: $X_i = \tanh(W_i * X_{i-1} + b_i)$;
- 3 Batch normalization: $N_1 = \frac{X_4 - \mu(X_4)}{\sigma(X_4) + \epsilon} \gamma + \beta$;
- 4 Denoised output: $O_{\text{NAM}} = N_2$;
- 5 Encode features: $E_{X_i} = f(W_{E_i} * E_{X_{i-1}} + b_{E_i})$;
- 6 Skip connections: $SCB_i = \text{ReLU}(W_{\text{SCB}_i} * E_{X_i} + b_{\text{SCB}_i}) + E_{X_i}$;
- 7 Decode features: $D_{z_i} = f(W_{D_i} * SCB + b_{D_i})$;
- 8 Dehazed output: $O_{\text{ACNM}} = \sigma(W_{\text{final}} * D_{X_2} + b_{\text{final}})$;
- 9 Extract features: $F_{X_i} = \text{ReLU}(W_{F_i} * F_{X_{i-1}} + b_{F_i})$;
- 10 Transform: $F_{X_4} = W_{F_4} * F_{X_3} + b_4$;
- 11 Concatenate: $O_{\text{FEM}} = \text{Concat}(I, F_{X_4})$;
- 12 Refinement: $D_{X_1} = \text{ReLU}(W_{D_1} * I + b_{D_1})$;
- 13 Further refine: $D_{X_2} = \tanh(W_{D_2} * D_{X_1} + b_{D_2})$;
- 14 Final transformation: $D_{X_3} = W_{D_3} * D_{X_2} + b_{D_3}$;
- 15 Concatenate: $O_{\text{DRM}} = \text{Concat}(I, D_{X_3})$;
- 16 Final output: $O_{\text{final}} = \alpha O_{\text{NAM}} + \beta O_{\text{ACNM}} + \gamma O_{\text{FEM}} + \delta O_{\text{DRM}}$;

To retain spatial and colour consistency, the output is concatenated with the original image:

$$O_{\text{FEM}} = \text{Concat}(I, F_{X_4}) \quad (20)$$

This fusion enables the model to blend learned features with original content, enhancing contrast and edge preservation.

3.5 | Final Output Integration

Outputs from all modules are aggregated through weighted element-wise summation:

$$O_{\text{final}} = \alpha O_{\text{NAM}} + \beta O_{\text{ACNM}} + \gamma O_{\text{FEM}} + \delta O_{\text{DRM}} \quad (21)$$

where α , β , γ , and δ are learnable weights optimized during training to balance contributions from each module. The overall procedure of the proposed work is shown in Algorithm 1.

3.6 | Loss Function

Throughout the training process, our model learns the function $F_w(I|G)$, where w represents the set of trainable parameters, I denotes the input image, and G represents the relevant, clear image. The mean square error (MSE) stands as the frequently employed loss function for image generation and is articulated

TABLE 2 | Impact of different loss combinations on image restoration performance.

Loss combination	PSNR (dB)	SSIM
$\mathcal{L}_{\text{MSE}}$ only	25.63	0.864
$\mathcal{L}_{\text{MSE}} + \mathcal{L}_{\text{SSIM}}$	26.72	0.879
$\mathcal{L}_{\text{MSE}} + \mathcal{L}_{\text{Perc}}$	26.15	0.872
$\mathcal{L}_{\text{MSE}} + \mathcal{L}_{\text{SSIM}} + \mathcal{L}_{\text{Perc}}$	27.31	0.889

as:

$$\mathcal{L}_{\text{MSE}} = \frac{1}{N} \sum_{x=1}^N \sum_{i=1}^3 \|\hat{y}_i(x) - y_i(x)\|^2 \quad (22)$$

where $\hat{y}_i(x)$ is the intensity of the i th colour channel of the pixel x in the input image, and $y_i(x)$ is the respective ground truth. MSE loss is useful for image restoration because of its simplicity, smoothness, and pixel-wise comparison.

However, this loss is sensitive to outliers and tends to excessively penalize the optimization process, which results in blurring in the reconstructed image. To mitigate this blurring, the proposed model incorporates a fusion of the other two loss functions: $\mathcal{L}_{\text{SSIM}}$ and $\mathcal{L}_{\text{Perc}}$, that are defined as

$$\mathcal{L}_{\text{SSIM}}(\hat{y}_i(x), y_i(x)) = 1 - \text{SSIM}(\hat{y}_i(x), y_i(x)) \quad (23)$$

SSIM is a perceptual metric used to quantify image deterioration by evaluating changes in structural information, luminance, and contrast. It aids in maintaining the structural integrity and perceptual quality of images. Perceptual loss is computed using features extracted from a pre-trained convolutional neural network, typically VGG19. This loss is the mean squared error among the feature representations of the ground truth and predicted images at a specific convolutional layer. The perceptual loss is defined as:

$$\mathcal{L}_{\text{Perc}}(\hat{y}_i(x), y_i(x)) = \frac{1}{N} \sum_{i=1}^N \|\phi(\hat{y}_i(x)) - \phi(y_i(x))\|^2 \quad (24)$$

where, ϕ is the feature extraction function from the VGG19 network. $\hat{y}_i(x)$ and $y_i(x)$ are the true and estimated output, respectively. N is the number of feature elements.

The combined loss function that includes MSE, SSIM loss, and perceptual loss can be defined as:

$$\text{Combined_loss} = \alpha \cdot \mathcal{L}_{\text{MSE}} + \beta \cdot \mathcal{L}_{\text{SSIM}} + \gamma \cdot \mathcal{L}_{\text{Perc}} \quad (25)$$

where α , β , and γ are hyperparameters that harmonize the various losses by assigning weights to each component. This ensures the model optimizes overall performance without emphasizing a single metric too heavily. The effectiveness and necessity of each loss component and its corresponding weight are validated via a sensitivity analysis, as detailed in Table 2.

3.6.1 | Effect of Loss Function Composition

We evaluated the necessity of each loss component by conducting an ablation study with different loss configurations. Table 2 shows that using all three loss functions—MSE, SSIM, and perceptual—leads to the best performance in terms of PSNR and SSIM and also promotes stable convergence during training. The values in Table 2 were obtained by training the model on the SOTS indoor dataset from RESIDE, using identical hyperparameters for fair comparison. The weight values for each loss were empirically selected to balance perceptual quality and reconstruction fidelity, based on the convergence behaviour during training.

4 | Experiments

In this section, we present a comprehensive evaluation of the proposed framework through a series of experiments. We begin by detailing the benchmark datasets and implementation configurations used for training and testing. Subsequently, we outline the baseline and recent state-of-the-art methods included for comparative analysis, along with the metrics adopted to assess image quality. This is followed by quantitative and qualitative results that illustrate the effectiveness of our method. Finally, we conduct an ablation study to examine the individual contribution of each module within the network.

4.1 | Implementation Details

4.1.1 | Datasets

We train our model using a combination of real-world and synthetic datasets. The I-HAZE and O-HAZE datasets, introduced by Ancuti et al. [38, 39], are real-world datasets captured using controlled haze machines in indoor and outdoor environments, respectively. I-HAZE consists of 35 pairs of high-resolution indoor images with and without haze, while O-HAZE contains 45 pairs of outdoor hazy images and their corresponding ground truths.

In addition, we utilize the RESIDE-6K dataset [40], a large-scale synthetic benchmark comprising 6000 paired indoor and outdoor images with varying haze densities. This diverse dataset helps improve the model's generalization across multiple scenarios. For further evaluation, we test the generalizability of our model on practical datasets, including SateHaze1k [41], a remote sensing dataset with three fog density levels (thin, moderate, and thick). Each level contains 320 training, 35 validation, and 45 test image pairs.

We also assess the method's performance on synthetic night-time scenes using selected samples from the Nighttime Hazy Middlebury (NHM) dataset [42], which provides hazy images with variable illumination and haze intensity. Given the similarity between haze in the atmosphere and scattering in underwater scenes, we extend our approach to underwater image enhancement. For this, we use the EUVP dataset [43], which includes 2185 real-world underwater images, with 2000 used for training and 185 for testing.

4.1.2 | Training Details

The method was implemented using the TensorFlow library, and every experiment was conducted on two NVIDIA GeForce RTX 4090 GPUs. We employed ADAM with default values for β_1 and β_2 for optimization, with a learning rate of 0.001. The proposed method was trained by optimizing the loss function $\mathcal{L}$ as defined in Equation (22). The network is designed to process input images of size 400×400 ; therefore, each image in the training dataset undergoes random cropping to maintain consistency and optimize performance.

We empirically set the loss weights as $\alpha = 1.0$, $\beta = 0.5$, and $\gamma = 0.2$ based on sensitivity analysis, aiming to balance perceptual quality and reconstruction fidelity while ensuring stable convergence. These values were selected by monitoring PSNR/SSIM trends and training stability on the validation set. To enhance model generalization and prevent overfitting, we applied random cropping, horizontal flipping, and brightness jittering during training.

4.1.3 | Evaluation Settings

To validate the proposed model, we contrast it with various advanced methods on synthetic, real-day, synthetic night vision, and remote sensing datasets. To further affirm the efficacy of the proposed algorithm, we conduct a comparative evaluation with underwater recovery algorithms using the EUVP dataset. The widely recognized full-reference image quality assessment metrics, mean-square error (MSE), peak signal-to-noise ratio (PSNR) and structural similarity index (SSIM) [44], are utilized to quantitatively evaluate the performance of the proposed approach and other dehazing methods. Specifically, PSNR measures the pixel-wise difference between the restored and ground-truth images, while SSIM assesses image similarity by considering luminance, contrast, and structural characteristics. Higher values of these metrics generally indicate superior performance. Additionally, the CIEDE2000 metric [45] is employed to quantify the colour discrepancy between the restored and ground-truth images. We report the average PSNR and SSIM of 26 state-of-the-art dehazing methods with different prior, supervised, semi-supervised, and unsupervised methods on the synthetic outdoor testing set (SOTS) dataset in Table 3. Since haze primarily impacts vision systems used outdoors, our evaluation mainly focuses on testing these methods using outdoor datasets. We have used O-HAZE and I-HAZE to evaluate the effectiveness of the proposed algorithm with the existing algorithms, including DCP [23], non-local [26], AOD-Net [11], GCAN [46], and MFAN [47].

Similarly, dehaze results on satellite remote sensing StateHaze1k datasets have been compared with DCP [23], MSBDN [48], ACER [49], TBN [50], and SGID [51]. To showcase the efficacy of the proposed model, we contrast it with several professional night-time dehazing methods on synthetic imaging datasets, including NDIM [52], GS [2], MRP [53], OSFD [42], and VDM [54].

Using the EUVP dataset, we evaluate the model performance by comparing it with several underwater restoration models, including PGC-DN [55], MSBDN [48], i-Dehaze [56], U-Shape [57], and WaveNet [58]. All evaluations on real and synthetic

hazy, night-vision, and underwater datasets have been done using PSNR and SSIM metrics.

4.2 | Comparison With State-of-the-Art Methods

4.2.1 | Qualitative Evaluation

In this section, we conduct a qualitative analysis comparing our approach with other leading dehazing methods, as illustrated in Figures 5–11. Figure 5 shows the visualization results from different methods in the synthetic test datasets, specifically SOTS-Outdoor [40].

Prior-based methods like DCP and non-local commonly exhibit excessive enhancement and loss of original colours. For instance, the sky regions may become excessively enhanced and are visible in all the images. AOD-Net either fails to eliminate haze in dense circumstances effectively or generates diminished luminosity compared to the original. In the second image, only GCAN and the proposed method effectively eliminate the haze from the motorcycle (bottom-left of the image). However, GCAN results in images with haze patches and distortions near the street light (visible upon closer inspection), while the proposed approach generates a more lifelike image.

We extended our qualitative evaluation to both synthetic and real-world datasets. For example, from the real-world datasets depicted in Figure 6, the proposed model diminishes more haze than the leading techniques in diverse environments. The outdoor dehazed images of DCP and non-locals exhibit severe colour distortion. The AOD-Net dehazed images show a noticeable haze residue. GCAN and MFAN display haze-distorted patches and perform well primarily on synthetic datasets such as RESIDE.

Figure 7 shows the dehazed results of the real-life I-Haze dataset. DCP shows noticeable dehazing effects in indoor images, whereas AOD-Net's images have visible haze left. GDN and MSBDN perform adequately well on synthetic datasets; however, their outcomes on real datasets are not as effective as ours. GDN's outcomes exhibit visible distortions, and MSBDN shows a significant colour deviation in the enlarged details. Our method outperforms others in both haze elimination and colour recovery.

To further validate real-world applicability, we employed samples from the Waterloo dataset and additional real-world images. Figure 8 shows a qualitative comparison of dehazing results on real-world images sourced from the Waterloo dataset and other real-world samples. The comparison includes DehazeNet [8], DHGAN [59], DehazeGAN [60], CNBID [14], and our proposed method. DehazeNet provides moderate haze suppression but leaves residual haze, particularly in the sky and distant objects. DHGAN and DehazeGAN improve dehazing depth but often generate unnatural colours or brightness, especially in facial and urban areas. CNBID enhances visibility yet tends to oversaturate and distort colour balance. In contrast, our method demonstrates superior haze removal while preserving realistic textures and tones across varied conditions—architecture, cityscapes, portraits, and foliage. These results verify the generalizability and robustness of our model to complex, real-world haze scenarios beyond synthetic datasets such as RESIDE and SOTS.

TABLE 3 | Quantitative comparisons on the SOTS dataset.

Method	Type	SOTS (outdoor)				SOTS (Indoor)		
		PSNR $\uparrow$	SSIM $\uparrow$	MSE $\downarrow$	CIEDE2000 $\downarrow$	PSNR $\uparrow$	SSIM $\uparrow$	MSE $\downarrow$
DCP [23]	Prior	18.54	0.710	0.259	10.199	18.87	0.794	0.190
BCCR [66]	Prior	15.71	0.769	—	—	—	—	—
NCP [26]	Prior	18.07	0.802	—	—	—	—	—
Dual-SaleNet [67]	Prior + supervised	21.76	0.909	—	—	—	—	—
DehazeNet [8]	Supervised	26.84	0.826	0.033	—	22.66	0.833	0.145
MSCNN [9]	Supervised	21.73	0.831	0.125	10.207	20.01	0.791	0.144
AOD-Net [11]	Supervised	24.08	0.873	0.127	7.287	21.01	0.816	0.118
GFN [27]	Supervised	21.67	0.852	0.040	—	22.44	0.884	0.069
GMAN [32]	Supervised	28.19	0.964	—	—	27.94	0.897	—
EPDN [68]	Supervised	22.57	0.863	—	—	—	—	—
GCANet [46]	Supervised	21.66	0.867	0.139	7.314	—	—	0.006
MSCNN-HE [69]	Supervised	22.72	0.871	—	—	—	—	—
FD-GAN [70]	Supervised	23.769	0.926	—	6.537	—	—	—
GFN-IJCV [71]	Supervised	24.21	0.849	—	—	—	—	—
MSFNet [72]	Supervised	30.07	0.939	—	—	—	—	—
Semi-Dehazing [73]	Semi-supervised	24.79	0.892	—	4.856	—	—	—
PSD [63]	Semi-supervised	20.49	0.844	—	14.292	—	—	—
SLAdhazing [74]	Self-supervised	24.33	0.932	—	—	—	—	—
CycleGAN [75]	Unsupervised	17.32	0.706	—	13.394	—	—	—
Cycle-Dehaze [76]	Unsupervised	18.60	0.797	—	13.967	—	—	—
Deep-DCP [77]	Unsupervised	20.99	0.893	—	—	—	—	—
YOLY [78]	Unsupervised	20.39	0.889	—	8.557	—	—	—
LIGHT-Net [79]	Unsupervised	23.11	0.917	—	—	—	—	—
Cycle-SNSPGAN [80]	Unsupervised	23.91	0.911	—	—	—	—	—
UCL-Dehaze [20]	Unsupervised	25.21	0.927	—	4.784	—	—	—
Ours	Supervised	28.44	0.967	0.025	4.150	28.42	0.899	0.057

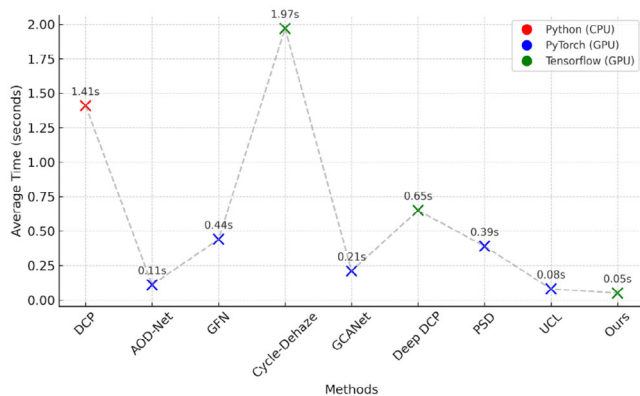
**FIGURE 5** | Visual results comparison on the SOTS-Outdoor dataset.

Figure 9 presents visual comparisons of various methods in the moderate Haze1K dataset. DCP exhibits severe colour distortion due to its strong reliance on atmospheric priors, leading to

inaccurate estimations in complex scenes. Although MSBDN and AECR produce better dehazing results, their restored images show noticeable discrepancies in colour and illumination compared to the original. TBN partially improves colour fidelity but compromises fine details in certain regions. SGID, on the other hand, introduces an unintended blue hue due to the incorporation of blur features within its network, affecting overall image quality. As depicted in the first image, our reconstructed image closely resembles the GT in overall colour and exhibits a more effective dehazing impact in the grass area. In general, our model demonstrates strong dehazing performance, producing images that closely match ground truth in terms of hue, brightness, detail, and clarity. This observation is further corroborated by quantitative analysis.

Moreover, we also perform evaluations on synthetic night-time images. Three typical synthetic examples with varying degrees of haze were chosen from the recently published “Nighttime Hazy Middlebury” (NHM) dataset, as illustrated in Figure 10. As shown in Figure 10, NDIM substantially enhances the visibility and

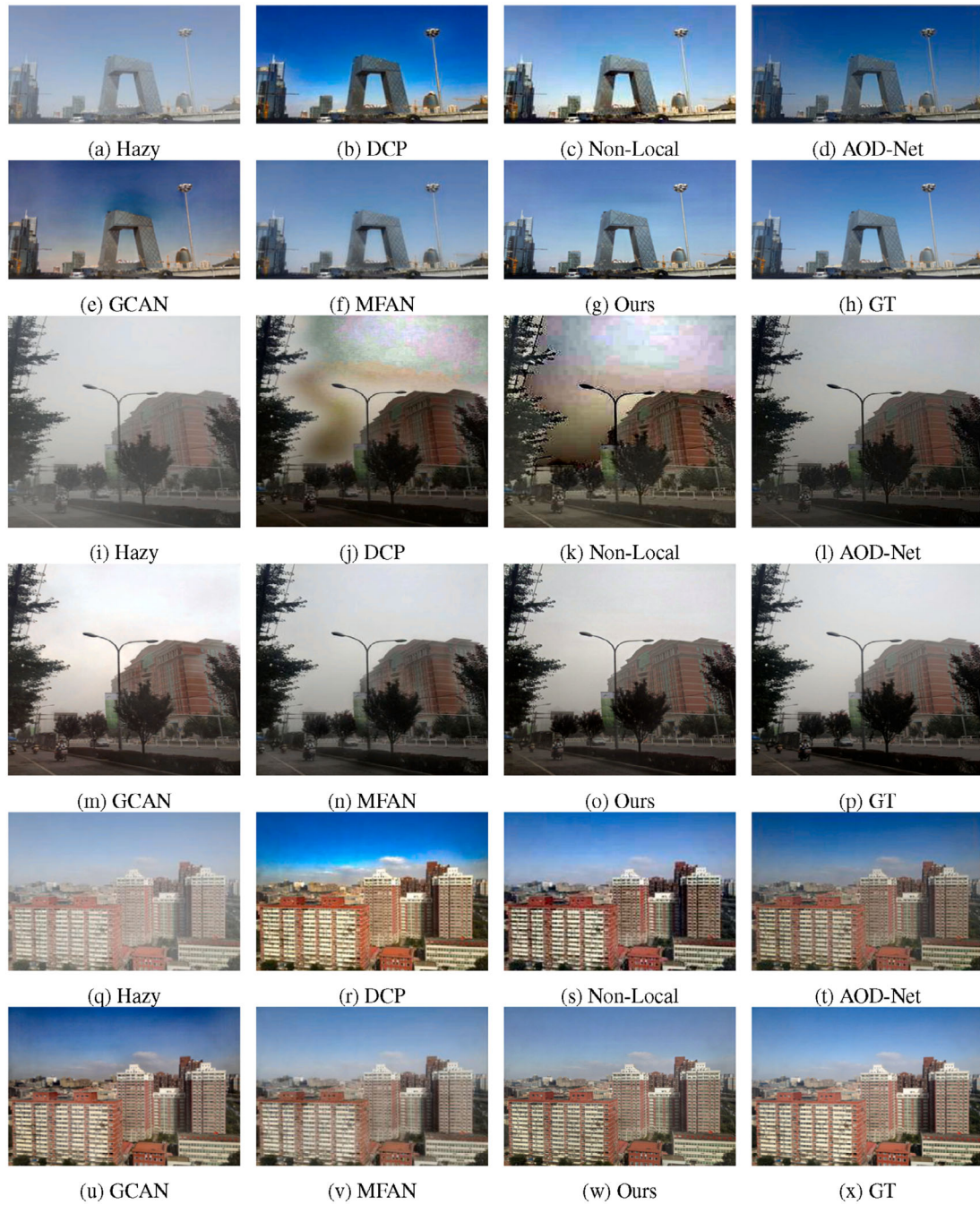


FIGURE 6 | Visual results comparison on the O-Haze dataset.

contrast of night-time hazy images. However, the dehazed results exhibit an unnatural appearance with uneven illumination. GS fails to compensate for brightness, resulting in dim images. MPRe effectively removes night-time haze but does not alleviate the glowing phenomena near light sources. OSFD does not recover details and all these methods amplify noise due to not addressing night-time hazy image noise, as seen in the second and last rows of Figure 10. VDM performs better in comparison to alternative methods; however, it fails to comply with the ground truth of these synthetic night-vision images. Overall, these algorithms fail to offer a holistic solution for night-time image dehazing. Conversely, the proposed framework enhances contrast, compensates for brightness, suppresses noise, and reveals details.

Light scattering and absorption by water lead to colour deviations in underwater images. Therefore, we utilized the model presented for underwater image restoration and contrasted it with various algorithms, including the most recent underwater restoration techniques such as U-shape and WaveNet. Figure 11 illustrates that our method outperforms others in the EUVP data set. In the same figure, the results of PGC-DN and MSBDN exhibit noticeable colour deviations. Although U-shape and WaveNet provide relatively better colour restoration, their overall image quality and texture preservation remain suboptimal. Specifically, in the last row, U-shape fails to retain fine texture details, while WaveNet produces darkened outputs. i-Dehaze delivers comparatively better performance but tends to overbrighten certain



FIGURE 7 | Visual results comparison on the I-Haze dataset.

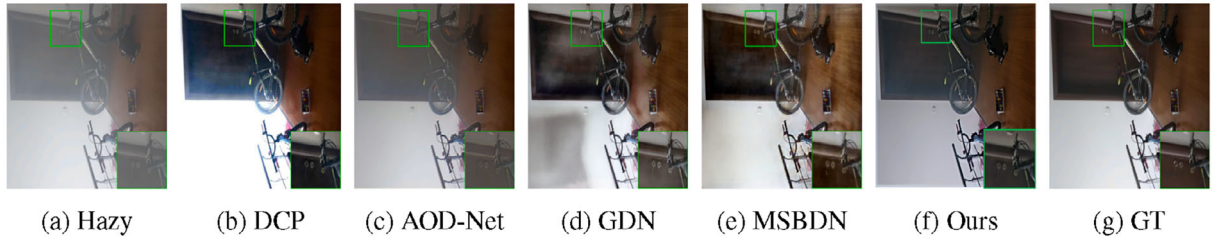


FIGURE 8 | Qualitative comparison of real-world hazy images from the Waterloo dataset and other real-world samples.

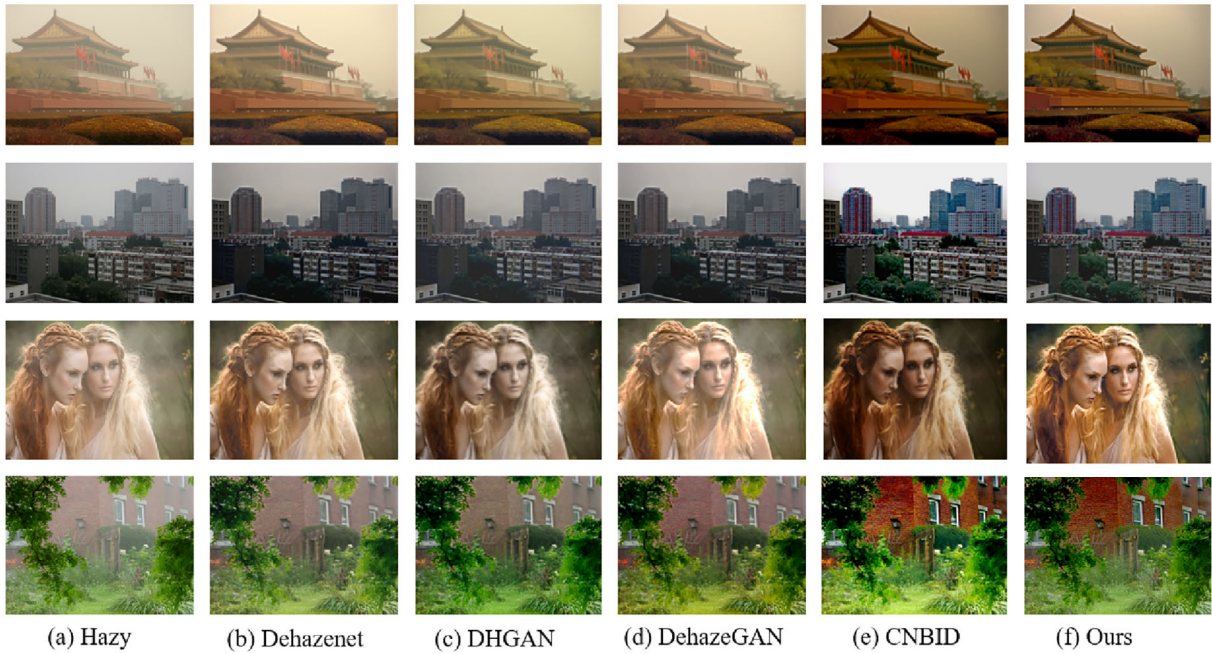


FIGURE 9 | Visual results comparison of the moderation on StateHaze1k dataset.

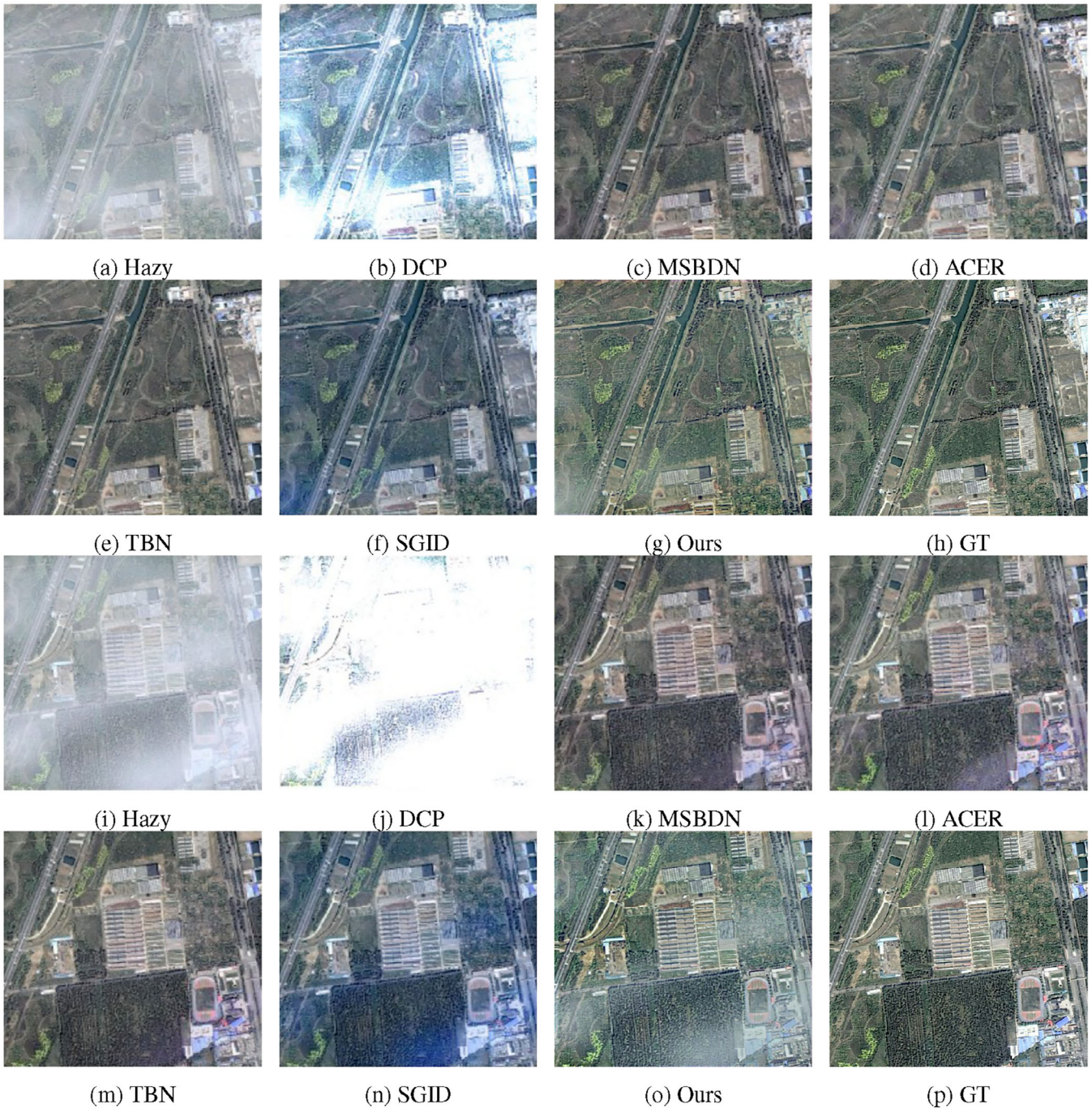


FIGURE 10 | Visual results comparison of night-time dehazing methods on synthetic scenes.

regions, leading to unnatural enhancements. In contrast, our method achieves superior visual quality by effectively balancing colour fidelity and detail preservation, which is particularly challenging in underwater environments due to light absorption and scattering effects. This robustness allows our approach to recover underwater scenes more realistically and comprehensively than existing methods.

However, while our method restores most regions effectively, minor colour distortions can still occur in distant background areas, as observed in the second image of Figure 11. These artefacts are attributed to depth-dependent scattering and intensity

attenuation in underwater images, which remain difficult to fully compensate for. Addressing this limitation will be part of our future work through the integration of depth-aware correction mechanisms to further enhance restoration performance in extreme depth zones.

4.2.2 | Quantitative Evaluation

Despite not incorporating the atmospheric scattering model, the proposed model exhibits remarkably strong dehazing capabilities. It surpasses numerous cutting-edge plug-in techniques, even

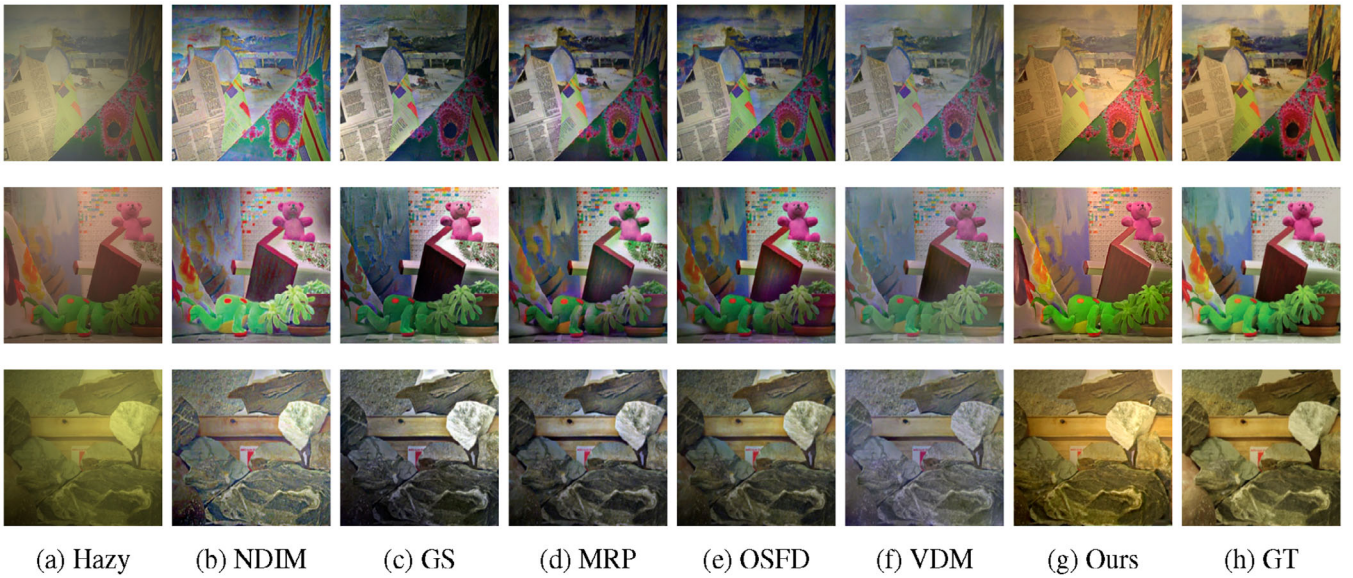


FIGURE 11 | Visual results comparison on the EUVP dataset.

on datasets generated with the atmospheric scattering model. As indicated in Table 3, our model delivers outstanding PSNR and SSIM performance, significantly surpassing other supervised dehazing algorithms in the outdoor SOTS dataset. Furthermore, compared to semi-supervised and unsupervised methods, our approach demonstrates unique advantages. It is possible that prior-based approaches typically emphasize improving overall image visibility while neglecting issues such as colour distortion and detail loss during restoration. For example, DCP struggles to dehaze sky regions effectively and often introduces artefacts. The proposed model effectively tackles many of these challenges and generates superior haze-free images.

We also assessed our model on the indoor SOTS dataset (see Table 3). In this setting, its effectiveness once again stands out significantly, surpassing other competing methods by a considerable gap in terms of PSNR and SSIM values. These results suggest that the proposed approach is not only viable but also effective in achieving remarkable dehazing outcomes for images captured both outdoors and indoors.

Table 4 presents the quantitative performance of various dehazing algorithms in real-world O-HAZE and I-HAZE datasets. On O-HAZE, our method achieves a PSNR of 24.92 and SSIM of 0.816, while on I-HAZE it reaches a PSNR of 26.18 and SSIM of 0.863. Compared to traditional and earlier deep learning approaches, the proposed model consistently shows high performance across both datasets. Notably, while CARL-Net and TUSR-Net show marginally higher PSNR on O-HAZE, our method achieves a better trade-off between contrast restoration and structural fidelity, reflected in a balanced SSIM score.

Recent state-of-the-art methods such as DNMGDT [22], which incorporates domain transfer and multi-prior guidance, exhibit moderate performance but fall short on both datasets compared to our method. Specifically, DNMGDT achieves a PSNR of 18.20 and SSIM of 0.793 on O-HAZE and 17.99/0.83 on I-HAZE. These

TABLE 4 | Quantitative comparisons on the O-HAZE and I-HAZE datasets.

Method	O-HAZE		I-HAZE	
	PSNR $\uparrow$	SSIM $\uparrow$	PSNR $\uparrow$	SSIM $\uparrow$
DCP [23]	16.78	0.653	14.43	0.725
AOD-Net [11]	17.56	0.650	13.98	0.732
GDN [10]	18.92	0.672	16.62	0.787
PGC-DN [55]	24.91	0.773	26.99	0.889
MSBDN [48]	24.36	0.749	23.93	0.891
CARL-Net [61]	25.83	0.807	25.43	0.880
TUSR-Net [62]	25.34	0.765	—	—
CAP [3]	14.55	0.567	13.42	0.650
NCP [26]	18.44	0.722	16.02	0.760
PSD [63]	16.45	0.610	14.49	0.626
DehazeFormer [12]	22.31	0.820	18.10	0.801
MGBL [64]	18.43	0.816	18.75	0.816
PFF [65]	22.93	0.750	18.41	0.782
GCA-Net [46]	18.05	0.743	18.57	0.744
UHD [17]	23.2	0.834	18.77	0.784
DNMGDT [22]	18.20	0.793	17.99	0.83
Ours	24.92	0.816	26.18	0.863

results underscore the superiority of our unified dehazing and detail enhancement strategy, particularly in preserving texture in dense haze conditions. Furthermore, Table 5 compares the performance on the StateHaze1k dataset, which includes a moderate haze density. Our method achieves a PSNR of 27.35 and an SSIM of 0.947, outperforming DehazeFormer [12] (26.53/0.941) and MFINEA [13] (27.61/0.942). These methods, despite employing

TABLE 5 | Quantitative comparison on StateHaze1k with moderate hazy density.

Method	StateHaze1k dataset	
	PSNR $\uparrow$	SSIM $\uparrow$
DCP [23]	9.78	0.591
MSBDN [48]	22.94	0.873
ACER [49]	24.03	0.879
TBN [50]	27.42	0.944
SGID [51]	23.95	0.935
Dehazeformer [12]	26.53	0.941
MFINEA [13]	27.61	0.942
Ours	27.35	0.947

transformer-based attention or multiscale feature fusion, do not exceed the balanced restoration performance achieved by our design.

Although certain methods like MSBDN and TBN occasionally report slightly higher individual PSNR or SSIM, our model provides a better global balance across all evaluated metrics, including CIEDE2000 and MSE. This consistency in performance across data sets and conditions highlights the adaptability and effectiveness of our modular design.

These comparisons collectively demonstrate that the proposed method not only competes strongly with but also exceeds recent state-of-the-art methods in most scenarios, validating its robustness for real-world image dehazing tasks.

To avoid subjective evaluation bias, we used two well-known full-reference metrics, PSNR and SSIM, for quantitative comparisons. Table 6 displays the evaluation results for Figure 10. Our method obtains the highest scores for PSNR and SSIM, indicating superior dehazing performance in hazy scenes at night. Furthermore, the NHM data set [42], comprising synthetic images 350 with various levels of haze, was used to assess the robustness of the proposed framework. Table 7 showcases the quantitative results on the EUVP dataset, while Figure 11 illustrates the qualitative results. As depicted in the table, our method surpasses the top underwater restoration algorithms, U-Shape and WaveNet, in terms of PSNR by 28.78, respectively, and SSIM by 0.885, respectively.

4.3 | Ablation Study

To assess the contribution of each module in our proposed strategy, we conducted an ablation study using a synthetic validation dataset. All the models are retrained in the same training strategy and evaluated on the outdoor synthetic validation set. The results, presented in Table 8, highlight the effect of removing individual modules on performance metrics, including mean squared error (MSE), peak signal-to-noise ratio (PSNR), and structural similarity index (SSIM).

4.3.1 | Effect of Different Components in Robust Convolutional Strategy

The NAM is designed to suppress noise in input data before further processing. Removing NAM led to a significant drop in performance, with MSE increasing from 0.005 to 0.012, PSNR decreasing from 25.480 dB to 19.860 dB, and SSIM dropping from 0.989 to 0.864. These results confirm that noise attenuation is crucial for enhancing the clarity of the input signal. The ACNM is responsible for adaptive feature extraction and learning spatial relationships. When ACNM was removed, MSE increased to 0.009, PSNR dropped to 22.609 dB, and SSIM decreased to 0.915. Although the degradation in performance is not as severe as the removal of NAM, these results highlight the module's role in feature adaptation. The DRM enhances fine details in the reconstructed signal. Removing DRM resulted in an MSE increase to 0.008, a drop in PSNR to 22.387 dB, and an SSIM reduction to 0.899. These results suggest that while DRM improves perceptual quality, its absence does not drastically degrade objective metrics. FEM is the most computationally intensive component in terms of performance. Without FEM, MSE increased to 0.007, PSNR dropped to 22.990 dB, and SSIM declined to 0.913. The results demonstrate that FEM plays a critical role in learning hierarchical representations, enabling the model to achieve state-of-the-art performance. The full model, integrating all modules, achieves the best results with MSE of 0.005, PSNR 25.480 dB, and SSIM of 0.989. The performance gains validate the necessity of each component, with FEM contributing the most to overall inference time.

These findings suggest that while removing individual modules may lead to computational savings, it often comes at the cost of decreased performance. FEM, although computationally expensive, is indispensable for maintaining high-quality outputs, whereas NAM, ACNM, and DRM collectively refine and optimize the results. The ablation study justifies the architecture's design choices, balancing accuracy with efficiency for practical deployment.

Visual analysis: Figure 12 further provides qualitative insight into the effect of each module through visual comparisons. Without NAM, the image retains visible haze, particularly in uniform regions, resulting in poor contrast and edge blurring—demonstrating the importance of early noise suppression. When DRM is removed, the output appears overly smoothed with diminished detail restoration, especially in textures and object boundaries, indicating its role in enhancing fine details. Excluding FEM results in moderate dehazing but compromises on structural depth and colour fidelity, particularly in sky and shadow areas. In contrast, the full model output is noticeably superior, displaying strong haze removal, colour consistency, and sharper edges. These visual results validate the necessity of each proposed module for achieving perceptually and quantitatively optimal dehazing performance.

4.4 | Efficiency Analysis

Efficiency is crucial for vision applications; therefore, we evaluated the performance of various advanced dehazing methods and listed their average run times in Figure 13. All methods were

TABLE 6 | Quantitative comparisons on synthetic night-time hazy images in Figure 10.

Image	NDIM [52]		GS [2]		MRP [53]		OSFD [42]		VDM [54]		Ours	
	PSNR $\uparrow$	SSIM $\uparrow$	PSNR $\uparrow$	SSIM $\uparrow$	PSNR $\uparrow$	SSIM $\uparrow$	PSNR $\uparrow$	SSIM $\uparrow$	PSNR $\uparrow$	SSIM $\uparrow$	PSNR $\uparrow$	SSIM $\uparrow$
Image 1	12.36	0.616	13.71	0.665	14.18	0.725	14.42	0.727	15.84	0.749	15.71	0.817
Image 2	14.69	0.683	14.92	0.715	14.86	0.740	14.37	0.748	16.66	0.789	18.42	0.798
Image 3	14.90	0.609	14.38	0.628	15.79	0.668	16.57	0.679	16.66	0.689	16.62	0.803

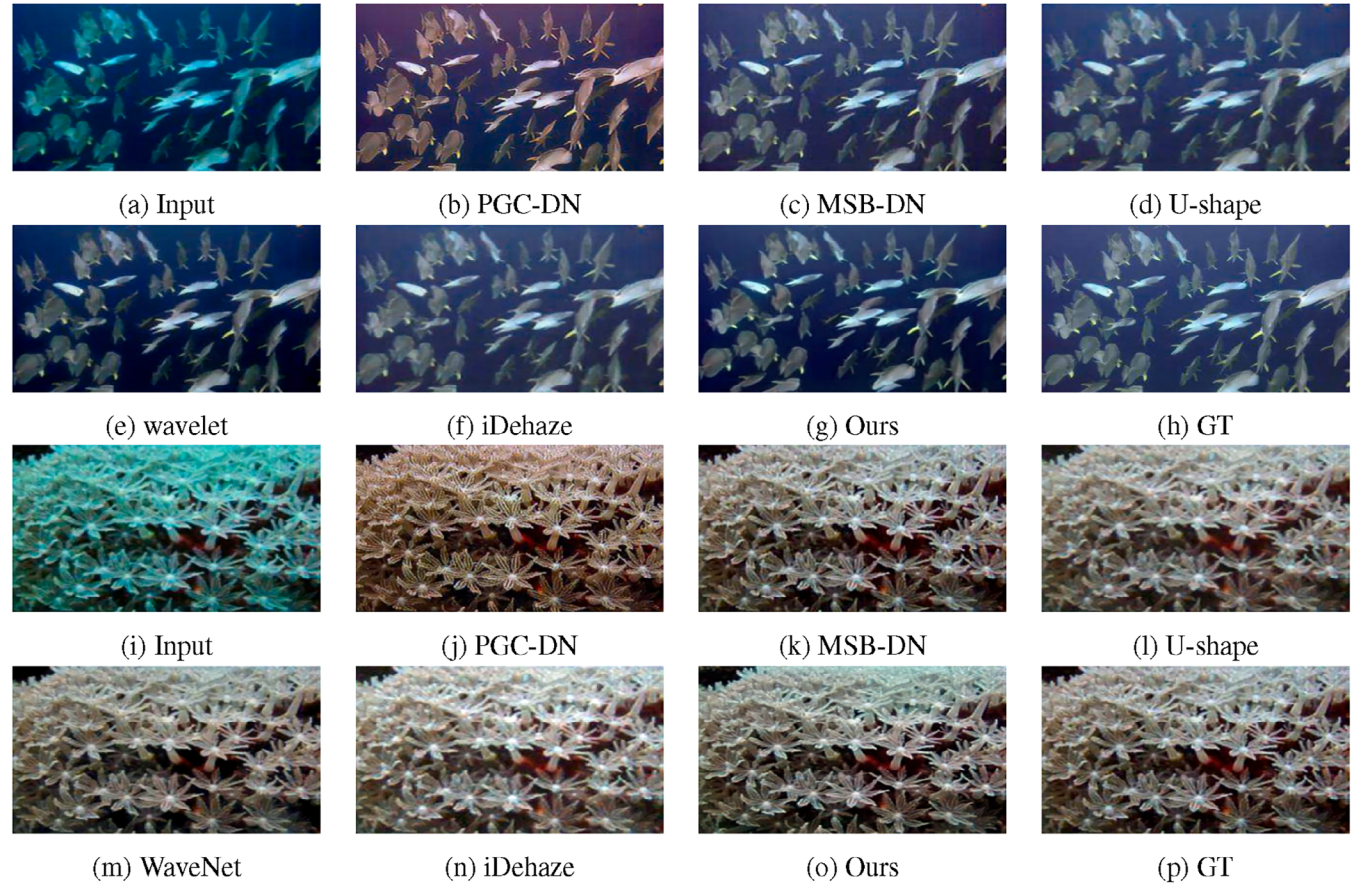
**FIGURE 12** | Visual results on a real-world hazy image under different architectural configurations of the proposed model.**FIGURE 13** | Average run time (seconds) of different methods tested on SOTS dataset.

TABLE 7 | Quantitative comparison on EUVP dataset.

Method	EUVP dataset	
	PSNR $\uparrow$	SSIM $\uparrow$
PGC-DN [55]	22.83	0.830
MSBDN [48]	27.91	0.866
I-Dehaze [56]	23.01	0.840
U-Shape [57]	28.75	0.882
WaveNet [58]	28.62	0.839
Ours	28.78	0.885

TABLE 8 | Ablation results on synthetic validation data. Each row shows the performance after removing one module from the full model.

Experiment	MSE $\downarrow$	PSNR $\uparrow$	SSIM $\uparrow$
w/o NAM	0.012	19.860	0.864
w/o ACNM	0.009	22.609	0.915
w/o DRM	0.008	22.387	0.899
w/o FEM	0.007	22.990	0.913
Full model	0.005	25.480	0.989

TABLE 9 | Model complexity and runtime comparison highlighting our method's efficiency.

Method	Parameters	FLOPs	Runtime (s)
FFDNet	486K	14.3G	0.062
DCPDN	18.6M	54.6G	0.134
GridDehazeNet	9.6M	36.1G	0.089
MSBDN	31.4M	125.8G	0.178
Ours	1.04M	6.47G	0.038

tested on a system with two NVIDIA GeForce RTX 4090 GPUs. The growing need for computing power places higher demands on hardware, limiting the widespread adoption of many models. In contrast, our approach does not require extensive computational resources. Our lightweight, efficient model processes a hazy image from the SOTS dataset in about 0.047 s, making it swifter and more effective than other dehazing algorithms.

Table 9 provides a quantitative assessment of computational complexity and inference speed for various dehazing methods. Our final model achieves the lowest parameter count at 1.04M and the minimum FLOPs of 6.47G, while also attaining the fastest runtime of 0.038 s. This clearly demonstrates the superiority of our approach in balancing efficiency and performance, making it highly suitable for real-time deployment and resource-constrained environments. The lightweight nature of our design significantly outperforms larger models such as MSBDN and DCPDN in both complexity and speed.

4.5 | Discussion

The modular design of the proposed dehazing network contributes to its flexibility and effectiveness in handling a wide range of haze conditions. By isolating tasks such as noise suppression, feature enhancement, contextual encoding, and detail refinement into dedicated components, the model is better equipped to preserve texture and contrast across complex scenes. A key strength lies in its ability to generalize across domains including night-time, underwater, and remote sensing imagery, without relying on domain-specific tuning or handcrafted priors. The architecture also supports fast inference, making it a practical choice for real-time and embedded applications. Despite its strengths, the current model is trained in a fully supervised setting, which may limit its applicability when paired training data are unavailable. Incorporating semi-supervised or unsupervised learning strategies in future work could further improve adaptability. Moreover, exploring temporal consistency may extend its use to video dehazing applications.

5 | Conclusions and Future Work

In this paper, we propose a lightweight convolutional neural network-based model to address the critical problems of image dehazing. The model consists of four primary components: (i) a noise attention module; (ii) an adaptive ConvNet module; (iii) a feature extraction module; and (iv) a detail refinement module. These modules collectively improve colour restoration, brightness, image detail, and feature extraction accuracy. Comprehensive experiments showcase the efficacy of these modules in overcoming common drawbacks of cutting-edge methods, including diminished colour richness and overly pronounced edges. Our model removes haze under diverse conditions, including synthetic night vision, remote sensing, and underwater image enhancement, producing more realistic output images.

As a supervised method, the proposed model requires large amounts of high-quality labelled training data, which may not always be available in real-world dehazing scenarios. Although it achieves strong performance across diverse conditions, this reliance on paired data can limit scalability when ground truth is scarce. Future work will therefore focus on extending the framework to semi-supervised and self-supervised paradigms, reducing dependency on paired datasets and improving adaptability to unseen environments. In addition, integrating the model as a preprocessing stage in object detection and tracking systems can enhance reliability in practical applications such as autonomous driving, surveillance, and remote sensing.

Author Contributions

Hira Khan: conceptualization, methodology, software, validation, formal analysis, original draft writing, review, and editing. **Sung Won Kim:** supervision, project administration, reviewing, and editing.

Acknowledgements

This research was supported in part by the Basic Science Research Program through the National Research Foundation of Korea (NRF),

funded by the Ministry of Education (NRF-2021R1A6A1A03039493) and in part by the NRF grant funded by the Korean government (MSIT) (NRF-2022R1A2C1004401).

Conflicts of Interest

The authors declare that they have no conflict of interest.

Data Availability Statement

Data sharing is not applicable to this article, as no data sets were generated during the current study.

References

- O. Elharrouss, N. Almaadeed, and S. Al-Maadeed, "A Review of Video Surveillance Systems," *Journal of Visual Communication and Image Representation* 77 (2021): 103116.
- L. Li, W. Feng, and J. Zhang, "Contrast Enhancement Based Single Image Dehazing VIA TV-LI Minimization," in *2014 IEEE International Conference on Multimedia and Expo (ICME)* (IEEE, 2014), 1–6.
- Q. Zhu, J. Mai, and L. Shao, "A Fast Single Image Haze Removal Algorithm Using Color Attenuation Prior," *IEEE Transactions on Image Processing* 24, no. 11 (2015): 3522–3533.
- N. Jiang, K. Hu, T. Zhang, W. Chen, Y. Xu, and T. Zhao, "Deep Hybrid Model for Single Image Dehazing and Detail Refinement," *Pattern Recognition* 136 (2023): 109227.
- H. Khan, M. Sharif, N. Bibi, et al., "Localization of Radiance Transformation for Image Dehazing in Wavelet Domain," *Neurocomputing* 381 (2020): 141–151.
- H. Khan, M. Sharif, N. Bibi, and N. Muhammad, "A Novel Algorithm for the Detection of Cerebral Aneurysm Using Sub-Band Morphological Operation," *The European Physical Journal Plus* 134, no. 1 (2019): 34.
- F. Yuan, Y. Zhou, X. Xia, X. Qian, and J. Huang, "A Confidence Prior for Image Dehazing," *Pattern Recognition* 119 (2021): 108076.
- B. Cai, X. Xu, K. Jia, C. Qing, and D. Tao, "DehazeNet: An End-to-End System for Single Image Haze Removal," *IEEE transactions on image processing* 25, no. 11 (2016): 5187–5198.
- W. Ren, S. Liu, H. Zhang, J. Pan, X. Cao, and M.-H. Yang, "Single Image Dehazing Via Multi-Scale Convolutional Neural Networks," in *Proceedings of the 14th European Conference Computer Vision-ECCV 2016* (Springer, 2016), 154–169.
- X. Liu, Y. Ma, Z. Shi, and J. Chen, "GridDehazeNet: Attention-Based Multi-Scale Network for Image Dehazing," in *Proceedings of the IEEE/CVF International Conference on Computer Vision* (IEEE, 2019), 7314–7323.
- B. Li, X. Peng, Z. Wang, J. Xu, and D. Feng, "AOD-Net: All-in-One Dehazing Network," in *Proceedings of the IEEE International Conference on Computer Vision* (IEEE, 2017), 4770–4778.
- Y. Song, Z. He, H. Qian, and X. Du, "Vision Transformers for Single Image Dehazing," *IEEE Transactions on Image Processing* 32 (2023): 1927–1941.
- H. Sun, B. Li, Z. Dan, et al., "Multi-Level Feature interaction and Efficient Non-Local information Enhanced Channel Attention for Image Dehazing," *Neural Networks* 163 (2023): 10–27.
- F. Alenezi, "Compared-Neighborhood Based Image Dehazing for Improved Visibility," *Engineering Applications of Artificial intelligence* 121 (2023): 106001.
- M. Wang, L. Liao, D. Huang, Z. Fan, J. Zhuang, and W. Zhang, "Frequency and Content Dual Stream Network for Image Dehazing," *Image and Vision Computing* 139 (2023): 104820.
- A. S. Parihar and A. Java, "Densely Connected Convolutional Transformer for Single Image Dehazing," *Journal of Visual Communication and Image Representation* 90 (2023): 103722.
- B. Xiao, Z. Zheng, Y. Zhuang, C. Lyu, and X. Jia, "Single UHD Image Dehazing via interpretable Pyramid Network," *Signal Processing* 214 (2024): 109225.
- A. Ali, R. Sarkar, and S. S. Chaudhuri, "Wavelet-Based Auto-Encoder for Simultaneous Haze and Rain Removal from Images," *Pattern Recognition* 150 (2024): 110370.
- Z. Chen, Z. He, and Z.-M. Lu, "DEA-Net: Single Image Dehazing Based On Detail-Enhanced Convolution and Content-Guided Attention," *IEEE Transactions on Image Processing* 33 (2024): 1002–1015.
- Y. Wang, X. Yan, F. L. Wang, et al., "Ucl-Dehaze: Towards Real-World Image Dehazing via Unsupervised Contrastive Learning," *IEEE Transactions on Image Processing* 33 (2024): 1361–1374.
- Z. Sun, Y. Li, S. Wu, and C. Zheng, "Dynamic-Routing 3D-Fusion Network for Remote Sensing Image Haze Removal," *IEEE Transactions on Geoscience and Remote Sensing* 63 (2025): 5606416.
- Y. Su, L. Shen, X. Liu, S. Zhou, and Y. Wang, "Real Scene Single Image Dehazing Network With Multi-Prior Guidance and Domain Transfer," *ISPRS Journal of Photogrammetry and Remote Sensing* 203 (2025): 128–140.
- K. He, J. Sun, and X. Tang, "Single Image Haze Removal Using Dark Channel Prior," *IEEE Transactions on Pattern Analysis and Machine Intelligence* 33, no. 12 (2010): 2341–2353.
- H. Khan, B. Xiao, W. Li, and N. Muhammad, "Recent Advancement in Haze Removal Approaches," *Multimedia Systems* 28, no. 3 (2022): 687–710.
- R. Fattal, "Dehazing Using Color-Lines," *ACM Transactions on Graphics* 34, no. 1 (2014): 1–14.
- D. Berman, S. Avidan, and S. Avidan, "Non-Local Image Dehazing," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (IEEE, 2016), 1674–1682.
- W. Ren, L. Ma, J. Zhang, et al., "Gated Fusion Network for Single Image Dehazing," in *Proceedings of the IEEE Conference on Computer vision and Pattern Recognition* (IEEE, 2018), 3253–3261.
- Y. Shao, L. Li, W. Ren, C. Gao, and N. Sang, "Domain Adaptation for Image Dehazing," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (IEEE, 2020), 2808–2817.
- A. Singh, A. Bhavne, and D. K. Prasad, "Single image dehazing for a variety of haze scenarios using back projected pyramid network," in *Proceedings of the European Conference on Computer Vision-ECCV 2020 Workshops* (Springer, 2020), 166–181.
- M. A. N. I. Fahim, N. Saqib, and H. Y. Jung, "Semi-Supervised Atmospheric Component Learning in Low-Light Image Problem," *PLOS ONE* 18, no. 3 (2023): e0282674.
- A. Ali, A. Ghosh, and S. S. Chaudhuri, "LIDN: A Novel Light Invariant Image Dehazing Network," *Engineering Applications of Artificial Intelligence* 126 (2023): 106830.
- Z. Liu, B. Xiao, M. Alrabeiah, K. Wang, and J. Chen, "Single Image Dehazing With a Generic Model-Agnostic Convolutional Neural Network," *IEEE Signal Processing Letters* 26, no. 6 (2019): 833–837.
- Q. Wang, Y. Yuan, J. Ren, and C. Li, "Underwater Image Enhancement Using Adaptive Retinex and Combined Priors," *Journal of Visual Communication and Image Representation* 71 (2020): 103008.
- A. Alenezi, R. Alshammari, and M. Alotaibi, "Real-Time Image Dehazing Algorithm Optimized for Low-Latency Environments," in *Proceedings of the International Conference on information Technology (ICIT)* (ACM, 2019) 3.
- Z. Lian, Y. Zhang, and Z. Wang, "Multi-Channel Feature Fusion Attention Dehazing Network," *PLOS ONE* 18, no. 7 (2023): e0286711.
- H. Qing, Y. Zhang, Y. Zhu, Y. Jiang, and M. Song, "Single Image Dehazing Algorithm based on Sky Segmentation and Optimal Transmission Maps," *The Visual Computer* 39, no. 3 (2023): 997–1013, <https://doi.org/10.1007/s00371-021-02380-3>.
- G. Sahu, A. Seal, D. Bhattacharjee, R. Frischer, and O. Krejcar, "A Novel Parameter Adaptive Dual Channel MSPCNN Based Single Image

- Dehazing for Intelligent Transportation Systems," *IEEE Transactions on Intelligent Transportation Systems* 24, no. 3 (2023): 3027–3047, <https://doi.org/10.1109/TITS.2022.3225797>.
38. C. Ancuti, C. O. Ancuti, R. Timofte, and C. De Vleeschouwer, "I-Haze: a Dehazing Benchmark With Real Hazy and Haze-Free Indoor Images," in *Proceedings of the Advanced Concepts for intelligent Vision Systems: 19th international Conference, ACIVS 2018* (Springer, 2018), 620–631.
39. C. O. Ancuti, C. Ancuti, R. Timofte, and C. De Vleeschouwer, "O-Haze: A Dehazing Benchmark With Real Hazy and Haze-Free Outdoor Images," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops* (IEEE, 2018), 754–762.
40. B. Li, W. Ren, D. Fu, et al., "Benchmarking Single-Image Dehazing and Beyond," *IEEE Transactions on Image Processing* 28, no. 1 (2018): 492–505.
41. B. Huang, L. Zhi, C. Yang, F. Sun, and Y. Song, "Single Satellite Optical Imagery Dehazing Using SAR Image Prior Based On Conditional Generative Adversarial Networks," in *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision* (IEEE, 2020), 1806–1813.
42. J. Zhang, Y. Cao, Z.-J. Zha, and D. Tao, "Nighttime Dehazing With A Synthetic Benchmark," in *Proceedings of the 28th ACM International Conference on Multimedia* (ACM, 2020), 2355–2363.
43. M. J. Islam, Y. Xia, and J. Sattar, "Fast Underwater Image Enhancement for Improved Visual Perception," *IEEE Robotics and Automation Letters* 5, no. 2 (2020): 3227–3234.
44. I. Q. Assessment, "From Error Visibility to Structural Similarity," *IEEE Transactions on Image Processing* 13, no. 4 (2004): 93.
45. G. Sharma, W. Wu, and E. N. Dalal, "The CIEDE2000 Color-Difference Formula: Implementation Notes, Supplementary Test Data, and Mathematical Observations," *Color Research & Application* 30, no. 1 (2005): 21–30.
46. D. Chen, M. He, Q. Fan, et al., "Gated Context Aggregation Network for Image Dehazing and Deraining," in *2019 IEEE Winter Conference on Applications of Computer Vision (WACV)* (IEEE, 2019), 1375–1383.
47. H.-H. Yang, C.-H. H. Yang, and Y.-C. J. Tsai, "Y-Net: Multi-Scale Feature Aggregation Network With Wavelet Structure Similarity Loss Function for Single Image Dehazing," in *ICASSP 2020-2020 IEEE international Conference on Acoustics, Speech and Signal Processing (ICASSP)* (IEEE, 2020), 2628–2632.
48. H. Dong, J. Pan, L. Xiang, et al., "Multi-Scale Boosted Dehazing Network With Dense Feature Fusion," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (IEEE, 2020), 2157–2167.
49. H. Wu, Y. Qu, S. Lin, et al., "Contrastive Learning for Compact Single Image Dehazing," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (2021), 10 551–10 560.
50. Y. Yu, H. Liu, M. Fu, J. Chen, X. Wang, and K. Wang, "A Two-Branch Neural Network for Non-Homogeneous Dehazing via Ensemble Learning," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (IEEE, 2021), 193–202.
51. H. Bai, J. Pan, X. Xiang, and J. Tang, "Self-Guided Image Dehazing Using Progressive Feature Fusion," *IEEE Transactions on Image Processing* 31 (2022): 1217–1229.
52. J. Zhang, Y. Cao, and Z. Wang, "Nighttime Haze Removal Based On A New Imaging Model," in *2014 IEEE International Conference on Image Processing (ICIP)* (IEEE, 2014), 4557–4561.
53. J. Zhang, Y. Cao, S. Fang, Y. Kang, and C. Wen Chen, "Fast Haze Removal for Nighttime Image Using Maximum Reflectance Prior," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (IEEE, 2017), 7418–7426.
54. Y. Liu, Z. Yan, A. Wu, T. Ye, and Y. Li, "Nighttime Image Dehazing Based on Variational Decomposition Model," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (IEEE, 2022), 640–649.
55. D. Zhao, L. Xu, L. Ma, J. Li, and Y. Yan, "Pyramid Global Context Network for Image Dehazing," *IEEE Transactions on Circuits and Systems for Video Technology* 31, no. 8 (2020): 3037–3050.
56. M. Mousavi, R. Estrada, and A. Ashok, "iDehaze: Supervised Underwater Image Enhancement and Dehazing via Physically Accurate Photorealistic Simulations," *Electronics* 12, no. 11 (2023): 2352.
57. L. Peng, C. Zhu, and L. Bian, "U-Shape Transformer for Underwater Image Enhancement," *IEEE Transactions on Image Processing* 32 (2023): 3066–3079.
58. P. Sharma, I. Bisht, and A. Sur, "Wavelength-Based Attributed Deep Neural Network for Underwater Image Restoration," *ACM Transactions on Multimedia Computing, Communications and Applications* 19, no. 1 (2023): 1–23.
59. W. Wu, J. Zhu, X. Su, and X. Zhang, "Dhgan: Generative Adversarial Network With Dark Channel Prior for Single-Image Dehazing," *Concurrency and Computation: Practice and Experience* 32, no. 18 (2020): e5263.
60. J. Zhu, L. Meng, W. Wu, D. Choi, and J. Ni, "Generative Adversarial Network-Based Atmospheric Scattering Model for Image Dehazing," *Digital Communications and Networks* 7, no. 2 (2021): 178–186.
61. D. Cheng, Y. Li, D. Zhang, N. Wang, X. Gao, and J. Sun, "Robust Single Image Dehazing Based on Consistent and Contrast-Assisted Reconstruction," *arXiv:2203.15325* (2022).
62. X. Song, D. Zhou, W. Li, et al., "TUSR-Net: Triple Unfolding Single Image Dehazing With Self-Regularization and Dual Feature to Pixel Attention," *IEEE Transactions on Image Processing* 32 (2023): 1231–1244.
63. Z. Chen, Y. Wang, Y. Yang, and D. Liu, "PSD: Principled Synthetic-to-Real Dehazing Guided By Physical Priors," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (IEEE, 2021), 7180–7189.
64. Z. Zheng, W. Ren, X. Cao, et al., "Ultra-High-Definition Image Dehazing via Multi-Guided Bilateral Learning," in *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (IEEE, 2021), 16180–16189.
65. K. Mei, A. Jiang, J. Li, and M. Wang, "Progressive Feature Fusion Network for Realistic Image Dehazing," in *Proceedings of the 14th Asian Conference on Computer Vision-ACCV 2018* (Springer, 2019), 203–215.
66. G. Meng, Y. Wang, J. Duan, S. Xiang, and C. Pan, "Efficient Image Dehazing With Boundary Constraint and Contextual Regularization," in *Proceedings of the IEEE International Conference on Computer Vision* (2013), 617–624.
67. Z. Li, C. Zheng, H. Shu, and S. Wu, "Dual-Scale Single Image Dehazing via Neural Augmentation," *IEEE Transactions on Image Processing* 31 (2022): 6213–6223.
68. Y. Qu, Y. Chen, J. Huang, and Y. Xie, "Enhanced Pix2Pix Dehazing Network," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (IEEE, 2019), 8160–8168.
69. W. Ren, J. Pan, H. Zhang, X. Cao, and M.-H. Yang, "Single Image Dehazing via Multi-Scale Convolutional Neural Networks With Holistic Edges," *International Journal of Computer Vision* 128 (2020): 240–259.
70. Y. Dong, Y. Liu, H. Zhang, S. Chen, and Y. Qiao, "FD-GAN: Generative Adversarial Networks With Fusion-Discriminator for Single Image Dehazing," in *Proceedings of the AAAI Conference on Artificial Intelligence* 34, no. 07 (2020): 10729–10736.
71. X. Zhang, H. Dong, Z. Hu, W.-S. Lai, F. Wang, and M.-H. Yang, "Gated Fusion Network for Degraded Image Super Resolution," *International Journal of Computer Vision* 128 (2020): 1699–1721.
72. X. Zhu, S. Li, Y. Gan, Y. Zhang, and B. Sun, "Multi-Stream Fusion Network With Generalized Smooth L1 Loss for Single Image Dehazing," *IEEE Transactions on Image Processing* 30 (2021): 7620–7635.
73. L. Li, Y. Dong, W. Ren, et al., "Semi-Supervised Image Dehazing," *IEEE Transactions on Image Processing* 29 (2019): 2766–2779.

74. Y. Liang, B. Wang, W. Zuo, J. Liu, and W. Ren, "Self-Supervised Learning and Adaptation for Single Image Dehazing," in *International Joint Conference on Artificial Intelligence* (Springer, 2022), 1137–1143.
75. J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired Image-to-Image Translation Using Cycle-Consistent Adversarial Networks," in *Proceedings of the IEEE International Conference on Computer Vision* (IEEE, 2017), 2223–2232.
76. D. Engin, A. Genç, and H. Kemal Ekenel, "Cycle-Dehaze: Enhanced CycleGAN for Single Image Dehazing," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops* (IEEE, 2018), 825–833.
77. A. Golts, D. Freedman, and M. Elad, "Unsupervised Single Image Dehazing Using Dark Channel Prior Loss," *IEEE Transactions on Image Processing* 29 (2019): 2692–2701.
78. B. Li, Y. Gou, S. Gu, J. Z. Liu, J. T. Zhou, and X. Peng, "You Only Look Yourself: Unsupervised and Untrained Single Image Dehazing Neural Network," *International Journal of Computer Vision* 129 (2021): 1754–1767.
79. A. Dudhane, P. W. Patil, and S. Murala, "An End-to-End Network for Image De-Hazing and Beyond," *IEEE Transactions on Emerging Topics in Computational Intelligence* 6, no. 1 (2020): 159–170.
80. Y. Wang, X. Yan, D. Guan, et al., "Cycle-SNSPGAN: Towards Real-World Image Dehazing via Cycle Spectral Normalized Soft Likelihood Estimation Patch GAN," *IEEE Transactions on intelligent Transportation Systems* 23, no. 11 (2022): 20368–20382.