# ReviveDiff: A Universal Diffusion Model for Restoring Images in Adverse Weather Conditions

Wenfeng Huang, Guoan Xu, Wenjing Jia, *Member, IEEE*, Stuart Perry,
and Guangwei Gao, *Senior Member, IEEE*

*Abstract*—Images captured in challenging environments–such as nighttime, smoke, rainy weather, and underwater–often suffer from significant degradation, resulting in a substantial loss of visual quality. The effective restoration of these degraded images is critical for the subsequent vision tasks. While many existing approaches have successfully incorporated specific priors for individual tasks, these tailored solutions limit their applicability to other degradations. In this work, we propose a universal network architecture, dubbed "ReviveDiff", which can address various degradations and restore images to their original quality by enhancing and restoring their details. Our approach is inspired by the observation that, unlike degradation caused by movement or electronic issues, quality degradation under adverse conditions primarily stems from natural media (such as fog, water, and low luminance), which generally preserves the original structures of objects. To restore the quality of such images, we leveraged the latest advancements in diffusion models and developed ReviveDiff to restore image quality from both macro and micro levels across some key factors determining image quality, such as sharpness, distortion, noise level, dynamic range, and color accuracy. We rigorously evaluated ReviveDiff on seven benchmark datasets covering five types of degrading conditions: Rainy, Underwater, Low-light, Smoke, and Nighttime Hazy. Our experimental results demonstrate that ReviveDiff outperforms the state-of-the-art methods both quantitatively and visually.

*Index Terms*—Image restoration, diffusion model, adverse conditions.

## I. INTRODUCTION

IMAGES captured in adverse environments, such as rain, fog, underwater conditions, and low-light scenarios, frequently endure significant degradation in quality. These natural factors disrupt light propagation, leading to substantial losses in visual clarity, color fidelity, and overall image visibility and usability. Furthermore, as these challenging conditions

often coincide with low illumination, the need for effective enhancement and restoration of image quality is paramount to ensuring that the subsequent vision tasks can be performed accurately and reliably.

Over the years, substantial progress has been made in addressing various image degradation challenges, including denoising [1], dedarkening [2], and deraining [3]. These methods typically rely on specific priors tailored to individual tasks. For instance, Retinex theory [4] has been successfully employed in many low-light enhancement techniques [2], while the atmospheric scattering model [5] has been commonly applied to guide the dehazing efforts. Fourier [6] and wavelet transformations have been extensively employed to address issues such as low resolution, blurring, and noise. However, these existing approaches are largely specialized, focusing on singular tasks with highly tailored solutions. This specialization, while effective within narrow scopes, limits its applicability across different types of degradation.

As research has progressed, more complex challenges, such as nighttime image dehazing [5], [7] and underwater image enhancement [8], [9], [10], [11], have emerged. Although these solutions have achieved impressive results by employing techniques such as Atmospheric Point Spread Function-guided Glow Rendering and posterior distribution processes for underwater color restoration, they remain constrained by their focus on specific degradation scenarios. Their tailored design potentially limits their applicability to other degradations.

Recognizing the limitations of task-specific models, recent efforts such as NAFNet [12], IR-SDE [13], and Restormer [14] have aimed to address common degradations with a single neural network architecture, dealing with issues such as blurring, noise, and low resolution. However, these methods often fall short in handling images from particularly challenging environments, such as those captured at night or underwater, which underscores the need for a universal framework capable of concurrently addressing diverse adverse weather conditions.

The root causes of image degradation vary significantly; for example, blurring is typically caused by motion, while noise and low resolution often stem from limitations in imaging equipment. In contrast, quality degradation under adverse conditions is predominantly due to natural phenomena such as fog, water, and low light. This type of degradation generally maintains the structural integrity of objects in the image, leading to less distortion and information loss than other degradation types. These characteristics have inspired us to
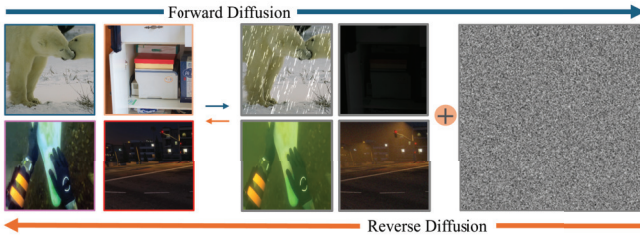
Fig. 1. Our ReviveDiff can address various adverse degradations (such as rain, low light, underwater, and nighttime dehazing) by enhancing and restoring their quality.

leverage general image attributes, such as edges, noise levels, and color distortions, to guide the distortion process across a wide range of adverse weather conditions, rather than relying on priors specific to an individual task.

To address this need for a more versatile and robust solution, we propose ReviveDiff, a groundbreaking universal diffusion model designed to restore image quality across a wide range of adverse environments. Our approach builds on the strengths of diffusion models, particularly their use of Stochastic Differential Equations (SDE) and Gaussian noise-based features, to create a model capable of handling an expansive feature space with exceptional learning capabilities. Unlike existing models, ReviveDiff is not confined to a single type of degradation but is equipped to tackle a wide range of adverse conditions and restore image qualities.

Specifically, what sets our ReviveDiff apart is its ability to address key factors that determine image quality, such as sharpness, contrast, noise, color accuracy, and distortion, from both a coarse-level overview and a detailed, fine-level perspective. Our model incorporates a novel multi-stage Multi-Attentional Feature Complementation module, which integrates spatial, channel, and point-wise attention mechanisms with dynamic weighting to balance macro and micro information. This design significantly enhances information integration, ensuring that the restored images are not only visually appealing but also retain the structural and color integrity necessary for subsequent tasks. Our comprehensive experiments, conducted across seven benchmark datasets and five types of adverse conditions–Rainy, Underwater, Low-light, Foggy, and Nighttime hazy–demonstrate that ReviveDiff consistently outperforms state-of-the-art (SOTA) methods in both quantitative measures and visual quality. This highlights the model's broad applicability and its potential to set new standards in image restoration under challenging conditions.

The main contributions of our work are listed as follows:

1) We developed ReviveDiff, a universal framework that can restore image quality across a wide range of adverse weather conditions, representing a significant advancement over the existing task-specific models.

2) We proposed a novel Coarse-to-Fine Learning Block (C2FBlock), which not only expands the receptive field through two distinct scale branches but also minimizes information loss. This design enables the network to capture complex and challenging degradations effectively.

3) We designed a novel Multi-Attentional Feature Complementation (MAFC) module that integrates spatial,

channel, and pixel attention mechanisms with dynamic weighting. This helps the model effectively complement information between macro and micro levels.

4) We introduced a unique prior-guided loss function that ensures optimal pixel restoration by leveraging edge information to refine object shapes and structures, while utilizing histogram information to guide accurate color correction under adverse conditions.

## II. RELATED WORK

In this section, we briefly review recent advancements in image restoration, particularly in challenging environments, and discuss the role of diffusion models in vision tasks.

### A. Image Restoration

Image restoration is a broad field aimed at recovering high-quality images from degraded inputs. Numerous single-task approaches have been proposed to tackle specific problems, such as deblurring, denoising, deraining, and super-resolution. For instance, MDARNet [15] employs dual attention for image deraining, and PREnet [16] focuses on progressive rain removal. In the context of dark conditions, Li et al. [17] introduced LLFormer with the Fourier Transform for low-light restoration. Underwater image restoration includes MLLE [18], which uses locally adaptive contrast, and UTrans [8], which integrates multi-scale feature fusion. For nighttime enhancement, Jin et al. [5] developed a gradient-adaptive convolution, and Yan et al. [7] used high-low frequency decomposition. Although these methods demonstrate strong performance, each primarily addresses a single type of degradation or environment, limiting their applicability across multiple adverse conditions.

In contrast, multi-task approaches aim to generalize or unify restoration across various tasks. SFNet [19] employed a multi-branch module with frequency information for broader image restoration, while Uformer [20] leveraged a U-shaped Transformer for multiple tasks. Restormer [14] was designed for high-resolution restoration, while MAXIM [21] introduced a multi-axis MLP structure to handle diverse degradations. Liu et al. [22] proposed PromptGIP, which adopts a visual prompting and question-answering paradigm to unify various image processing tasks. Zheng et al. [23] presented a diffusion-based universal restoration model with selective hourglass mapping, and Luo et al. [24] combined vision-language models for multi-task image restoration. Liu et al. [25] introduced a lightweight Task-Plugin module featuring a dual-branch architecture within a diffusion model to supply task-specific priors for low-level tasks. Sun et al. [26] proposed Histoformer, with a dubbed histogram self-attention mechanism for image restoration. By accommodating different degradation types, these multi-task solutions offer broader applicability, though balancing performance across diverse conditions remains an ongoing challenge.

### B. Diffusion Models in Vision

The diffusion models [27] have emerged as powerful tools in the field of image generation and restoration and have achieved
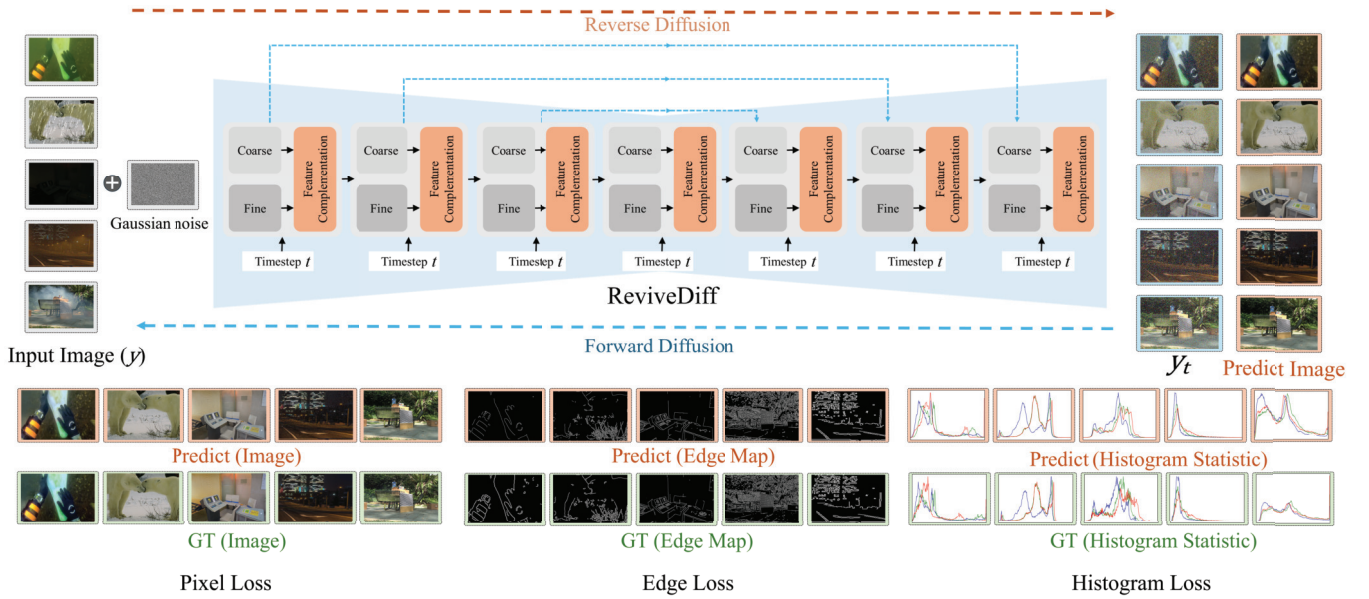
Fig. 2. The basic framework of our proposed ReviveDiff, which is a U-shaped latent diffusion model with stacked Coarse-to-Fine Learning Blocks (C2FBlock) and Multi-Attentional Feature Complementation (MAFC) modules, guided by a combined loss to tackle diverse real-world image restoration tasks.

significant improvement. These models work by degrading a signal through Gaussian noise and then restoring it through a reverse process. Fei et al. [28] introduced a unified image restoration model that utilizes a diffusion prior. In another significant contribution, Luo et al. proposed IR-SDE [13], which focuses on image restoration through mean-reverting stochastic differential equations.

The application of diffusion models has also extended to enhancing images captured in adverse conditions. For example, Jiang et al. proposed LightenDiffusion [29], a latent-Retinex-based diffusion model specially designed for low-light image enhancement. For underwater image enhancement, Du et al. proposed UIEDP [11], which employs a diffusion prior tailored for underwater images, and Tang et al. [30] introduced a transformer-based diffusion model. Zheng et al. [23] proposed a universal image restoration framework based on a diffusion model that leverages a selective hourglass mapping strategy. Wang et al. [31] introduced RDMD, a framework that leverages a single pre-trained diffusion model to derive two complementary regularizers. Kim et al. [32] proposed region-aware diffusion models for image inpainting. DIFFPIR [25], a lightweight task-plugin module introduced by Liu et al., features a dual-branch architecture within a diffusion model that provides task-specific priors for low-level tasks. The flexibility of diffusion models in accommodating different scenarios in these areas has inspired us to leverage the powerful generalization capabilities of diffusion models to build a universal solution–ReviveDiff– that can effectively address a wide range of adverse degradations. Furthermore, to further enhance its effectiveness in tackling diverse real-world image restoration tasks, we adopted the Mean-Reverting Stochastic Differential Equations (SDEs) based diffusion models [13], [33], which implicitly model the degradation and apply to diverse tasks without changing the architecture.

## III. METHODOLOGY

As illustrated in Fig. 2, our ReviveDiff is a U-shaped latent diffusion model, with stacked Coarse-to-Fine Blocks (C2FBlocks) and Multi-Attentional Feature Complementation (MAFC) modules to tackle diverse real-world image restoration tasks. Compared with other diffusion models used for image enhancement, our ReviveDiff performs fusion at both low-resolution, macro-level latent space and high-resolution, micro-level latent space from the original input for the decoding process, and under the guidance of fine granularity.

Building on the insights from the previous research [12], [13], [33], the C2FBlock introduces a dual-branch structure purposefully designed to capture features at varying levels of granularity. Specifically, as shown in Fig. 3, the Coarse Branch allows the fusion to be performed with a larger receptive field ($31 \times 31$), enabling it to capture broader contextual information, while the Fine Branch utilizes a focused receptive field of $3 \times 3$ to capture finer, more detailed features.

Then, to effectively integrate these coarse and fine features, we developed a Multi-Attentional Feature Complementation module (see Fig. 4), which incorporates three distinct types of attention mechanisms and employs dynamic weighting to adjust the balance between the contributions of Coarse and Fine features. This ensures an optimal balance that enhances the model's ability to restore image quality across diverse scenarios accurately.

### A. Mean-Reverting Stochastic Differential Processes

In this work, we leverage a probabilistic diffusion approach to enhance visibility in low-light images. Specifically, our approach is based on a score-based generative framework that utilizes Mean-Reverting SDE [13] as the base diffusion framework to model the image-reviving diffusion process. Fig. 2 illustrates this process. The forward SDE gradually

transforms the initial data distribution $y_0$ into a Gaussian noise representation $y_T$ over $T$ steps. The objective of the reverse SDE process is then to reconstruct a high-quality image from this noisy representation $y_T$.

*1) Forward SDE:* The forward SDE process is defined as the transformation of an input high-quality image into Gaussian noise. Mathematically, this process is expressed as

$$dy = \alpha(t)(v - y)\,dt + \beta(t)dV, \qquad (1)$$

where $\alpha(t)$ regulates the speed of mean reversion, $\beta(t)$ denotes the level of stochastic volatility, and $V$ is the standard Wiener process.

To achieve a solvable form of the forward SDE, we set the condition $\frac{\beta(t)^2}{\alpha(t)} = 2\kappa^2$, with $\kappa^2$ symbolizing the stationary variance for any $t \in [0, T]$. Assuming $y(s)$ depicts the noisy image derived from high-quality input at time $s < t$, the solution to Eq. 1 is given by [13]:

$$y(t) = v + (y(s) - v)e^{-\hat{\alpha}(s:t)} + \int_s^t \beta(u)e^{-\hat{\alpha}(u:t)}dV(u), \qquad (2)$$

where $\hat{\alpha}(s:t) = \int_s^t \alpha(u)du$. The distribution of $y_t$ is determined as follows:

$$y_t \sim p_t(y) = \mathcal{N}(y_t|\mu_t(y), \sigma_t^2), \qquad (3)$$

where

$$\mu_t(y) = \mu + (y(s) - \mu)e^{-\hat{\alpha}(s:t)}, \qquad (4)$$

and

$$\sigma_t^2 = \kappa^2 \left(1 - e^{-2\hat{\alpha}(s:t)}\right). \qquad (5)$$

*2) Reverse SDE:* The goal of the Reverse SDE process is to reconstruct the high-quality image from its degraded representation $y(T)$. As established in earlier research, the reverse SDE is formulated as

$$dy = \left[\alpha(t)(v - y) - \beta(t)^2 \nabla_y \log p_t(y)\right] dt + \beta(t)d\tilde{V}, \qquad (6)$$

where $\tilde{V}$ signifies another instance of the Wiener process.

The training mechanism utilizes high-quality images to enable the calculation of the gradient score:

$$\nabla_y \log p_t(y) = -\frac{y(t) - \mu_t(y)}{\sigma_t^2}, \qquad (7)$$

facilitating the sampling of $y(t)$ as $y(t) = \mu_t(y) + \sigma_t \zeta_t$, with $\zeta_t \sim \mathcal{N}(0, I)$ indicating Gaussian noise.

Thus, the score function is derived as

$$\nabla_y \log p_t(y) = -\frac{\zeta_t}{\sigma_t}. \qquad (8)$$

### B. Coarse-to-Fine Learning

Unlike the existing diffusion models used for image enhancement, our ReviveDiff performs feature fusion in the latent space at both macro and micro levels and fuses them under the guidance of fine granularity. This is achieved through a series of stacked C2FBlocks at the U-Net architecture's core.

As illustrated in Fig. 3, our C2FBlock is specially designed to capture features at two distinct levels of granularity: The Fine Branch is responsible for capturing the image's detailed, localized features with a small receptive field, whereas the Coarse Branch operates on a significantly larger receptive field,
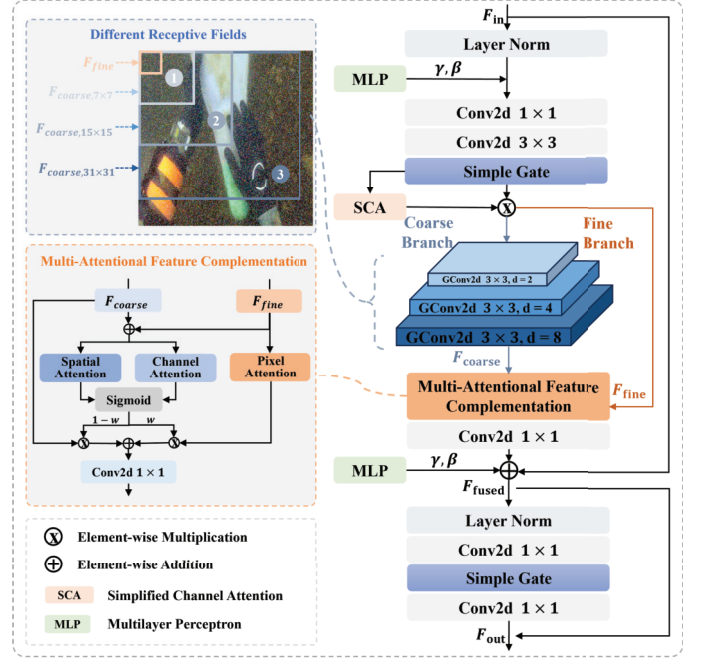


Fig. 3. The Coarse-to-Fine Learning Block (C2FBlock) at the core of our proposed ReviveDiff network.

focusing on the image's global contextual information in low-resolution space. This dual-branch approach ensures that both macro-level contextual information and detailed, pixel-level details are effectively captured and integrated.

In addition, to improve computational efficiency, we adopt the Simple Gate (SC) activation [12] to replace more complex nonlinear activation functions. The Simple Gate achieves the effect of nonlinear activation using a single multiplication operation, which is especially advantageous in preserving critical information in regions of low pixel values.

Specifically, let $\mathcal{F}_{in}$ represent the input feature. The features extracted by the Fine Branch can be represented as

$$\mathcal{F}'_{fine} = DWConv(LayerNorm(\alpha_1 \odot \mathcal{F}_{in} + \beta 1)). \qquad (9)$$

The Fine Branch uses a $3 \times 3$ receptive field to capture detailed local features, producing $\mathcal{F}'_{fine}$. Then, the final output of the Fine Branch $\mathcal{F}_{fine}$ is computed using SimpleGate as $SG$ and a Simplified Channel Attention as $SCA$, as

$$\mathcal{F}_{fine} = SCA(SG(\mathcal{F}'_{fine})). \qquad (10)$$

In contrast to the Fine Branch, the Coarse Branch focuses on global information and operates on a significantly larger receptive field of $31 \times 31$, as depicted in Fig. 4. However, existing approaches, such as the non-local mechanism, vision transformer blocks, and convolutional layers with large kernel sizes, tend to be computationally expensive, consuming significant resources and slowing down processing speeds. To address this issue, we utilize grouped-dilated convolutions to efficiently expand the receptive field size without substantially increasing the computation load.

To mitigate the information loss due to the gridding effect commonly associated with dilated convolutions, we employ a stacking strategy rather than a single convolutional layer.
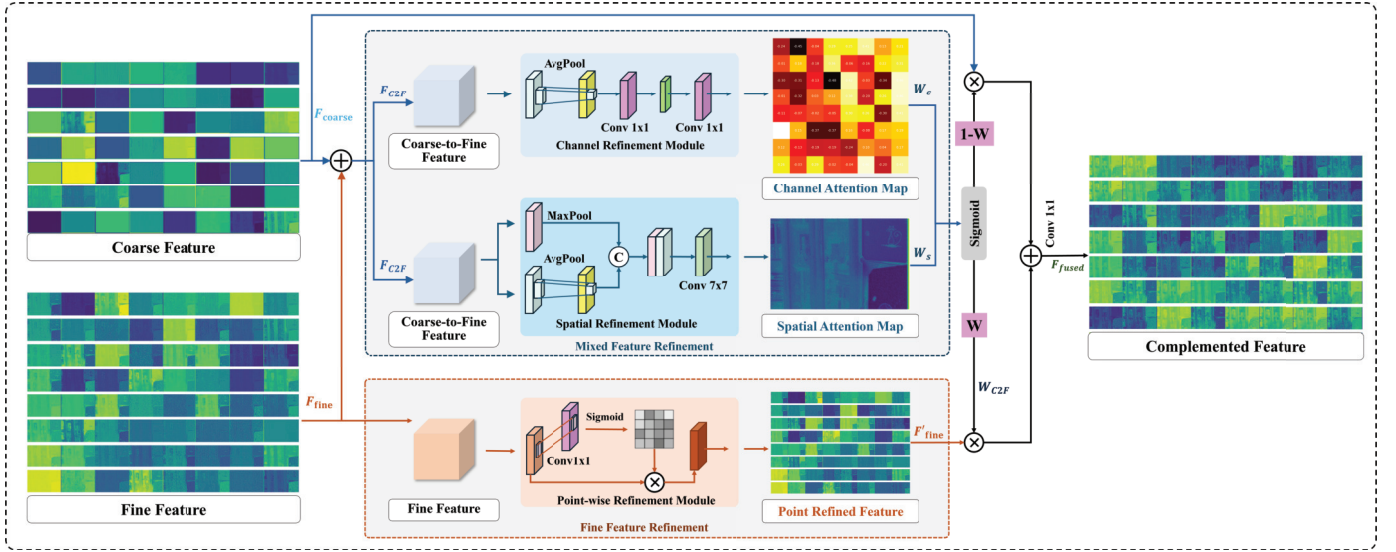
Fig. 4. Our Multi-Attentional Feature Complementation (MAFC) module adaptively integrates coarse and fine features through complementary refinement pathways. The coarse stream undergoes dual-domain enhancement via channel-wise recalibration and spatial context aggregation, while a point-wise feature refinement module refines the fine branch. The two branches are then fused through an adaptive weighting strategy, resulting in complemented feature maps that preserve both global context and local details.

Specifically, we apply increasing dilation rates of 2-4-8 to achieve a large receptive field while minimizing the loss of fine-grained details. Finally, grouped convolutions are used to prevent a substantial increase in parameters, ensuring that the model remains computationally efficient.

Thus, denoting grouped-dilated convolution with a dilation rate of $i$ as $f_{r=i}$, the Coarse Feature, $\mathcal{F}_{coarse,size}$, where $size$ represents its kernel size, can be obtained as

$$
\begin{cases}
\mathcal{F}_{coarse,7\times7} = f_{r=2}(\mathcal{F}_{fine}) \\
\mathcal{F}_{coarse,15\times15} = f_{r=4}(\mathcal{F}_{coarse,7\times7}) \\
\mathcal{F}_{coarse,31\times31} = f_{r=8}(\mathcal{F}_{coarse,15\times15}).
\end{cases}
\tag{11}
$$

This coarse-to-fine learning process enables our model to capture both broad contextual information and focused local details, ensuring a balanced and thorough restoration of images across different types of degradations. As experimental results show, by fusing coarse and fine features in the latent space, our ReviveDiff achieves superior performance in a range of challenging image enhancement tasks.

### C. Multi-Attentional Feature Complementation

As illustrated in Fig. 4, the proposed Multi-Attentional Feature Complementation (MAFC) module adaptively integrates coarse and fine features via complementary refinement branches. Specifically, coarse features undergo channel and spatial refinements, while fine features are refined by a point-wise feature refinement module. The two branches are then fused through an adaptive weighting strategy, resulting in complemented feature maps that preserve both global context and local details.

Given the Fine Features $\mathcal{F}_{fine}$ and the Coarse Features $\mathcal{F}_{coarse}$ obtained from Coarse-to-Fine Learning, the significant granularity gap between them poses a challenge for effective fusion and refinement.

To address this issue, we propose a Multi-Attentional Feature Complementation (MAFC) module, which utilizes attention-based mechanisms [39], [40] to dynamically compute feature weights both spatially [41] and across channels [42]. This allows for realignment and complementary fusion of the Coarse and Fine features, ensuring that the resulting feature map maintains critical information from both levels of granularity adaptively and complementarily.

For spatial dimension alignment, we utilize a spatial attention mechanism to learn the spatial weight map $W_s \in \mathbb{R}^{H\times W}$. This map highlights the importance of different regions within an image. Meanwhile, the channel weight is computed as $W_c \in \mathbb{R}^{C\times1\times1}$, indicating the importance of each channel for a given task.

Denote $\mathcal{F}_{C2F} = \mathcal{F}_{coarse} + \mathcal{F}_{fine}$ as the feature map integrating both Coarse and Fine Features, the spatial weight $W_s$ and the channel weight $W_c$ can be calculated as

$$
\begin{cases}
W_s = Cov_{7\times7}([GAP_s(\mathcal{F}_{C2F}), GMP_s(\mathcal{F}_{C2F})]), \\
W_c = Conv_{1\times1}(ReLU(Conv_{1\times1}(GAP_c(\mathcal{F}_{C2F})))).
\end{cases}
\tag{12}
$$

Here, $GAP$ refers to Global Average Pooling, $GMP$ denotes Global Max Pooling, and $s$ and $c$ indicate that the GAP/GMP operation is conducted along the spatial or channel dimensions, respectively.

Thus, the dynamic combination weight $W_{C2F}$ recalibrating $\mathcal{F}_{coarse}$ and $\mathcal{F}_{fine}$ can be defined as

$$
W_{C2F} = \sigma(W_s + W_c),
\tag{13}
$$

where $\sigma$ denotes the Sigmoid function.

Furthermore, to further refine the fine features $\mathcal{F}_{fine}$, we employ a Pixel Attention [43] scheme to establish attention coefficients at the pixel level, aiming to obtain a more sophisticated feature map $\mathcal{F}'_{fine}$:

$$
\mathcal{F}'_{fine} = \sigma(Conv_{1\times1}(\mathcal{F}_{fine})) \cdot \mathcal{F}_{fine}.
\tag{14}
$$

Fig. 5. Visual comparison of the deraining results obtained with our ReviveDiff and SOTA approaches on Rain100L [38]. The small images enclosed in red below each result provide an enlarged view of the areas highlighted by red rectangles.

Finally, the feature map $\mathcal{F}_{fused}$ can be generated by weighted recalibration as

$$\mathcal{F}_{fused} = Conv_{1\times 1}(\mathcal{F}'_{fine} \cdot W_{C2F} + \mathcal{F}_{coarse} \cdot (1 - W_{C2F})). \quad (15)$$

Then, the fused feature $\mathcal{F}_{fused}$ is processed with Layer Normalization (LN) and subsequently modified by the scale and shift factors $\alpha_2$ and $\beta_2$ as

$$\mathcal{F}'_{out} = LM(\alpha_2 \odot \mathcal{F}_{fused} + \beta_2). \quad (16)$$

Finally, the output feature of the C2FBlock, denoted as $\mathcal{F}_{out}$, is obtained as

$$\mathcal{F}_{out} = Conv_{1\times 1}(SG(Conv_{1\times 1}(\mathcal{F}'_{out}))). \quad (17)$$

This multi-attentional feature complementation allows for highly effective integration of Coarse and Fine features, ensuring that both global context and fine details are preserved and enhanced. This results in a feature map that is not only contextually aware but also rich in local details, providing the foundation for superior image restoration performance.

### D. Prior-Guided Loss Functions

In many related research fields, pixel-based loss functions such as L1, MSE, PSNR [63], and SSIM [64] Losses are widely used. These loss functions aim to minimize pixel-wise differences between the generated image $I_{gen}$ and the ground truth image $I_{gt}$. The Pixel-based Loss, $L_{pixel}$, is defined as

$$L_{pixel}(I_{gen}, I_{gt}) = \frac{1}{N}\sum_{i=1}^{N}|I_{gen,i} - I_{gt,i}|, \quad (18)$$

where $N$ denotes the total number of pixels in the image.

However, for low-quality images captured from adverse natural conditions, visibility is predominantly affected by the medium of light propagation, such as fog or underwater environments, rather than by distortion or blur. As a result,

objects within these images tend to retain their regular edges, shapes, and structures, even though overall visibility may be compromised. Leveraging this observation, incorporating edge and color information into the loss function can further enhance the training process by ensuring that key structural and color features are preserved during restoration and thus be advantageous.

*1) Edge Loss:* In our study, to preserve structural information, we introduce an Edge Loss that focuses on maintaining the integrity of edges in the enhanced image. We utilize the Canny edge detector to extract the edge information from both the enhanced image $I_{\text{gen}}$ and the corresponding reference image $I_{\text{gt}}$. The Edge Loss, $L_{\text{edge}}$, is defined as

$$
\begin{aligned}
&L_{\text{edge}}(I_{\text{gen}}, I_{\text{gt}}) \\
&= \left\| \mathcal{E}(I_{gen}; \sigma, T_{\text{low}}, T_{\text{high}}) - \mathcal{E}(I_{gt}; \sigma, T_{\text{low}}, T_{\text{high}}) \right\|_1 .
\end{aligned} \tag{19}
$$

Here, $\mathcal{E}(\cdot; \sigma, T_{\text{low}}, T_{\text{high}})$ denotes the edge map obtained by applying the Canny edge detector with a Gaussian smoothing parameter $\sigma$ and hysteresis thresholding parameters $T_{\text{low}}$ and $T_{\text{high}}$. The Canny edge detection process, $\mathcal{E}$, includes gradient magnitude computation using a Gaussian filter with standard deviation $\sigma$, non-maximum suppression, and edge tracking by hysteresis using thresholds $T_{\text{low}}$ and $T_{\text{high}}$. The $L_1$ norm $\|\cdot\|_1$ calculates the sum of absolute differences between the edge maps of the enhanced and reference images, ensuring the preservation of edge information during image enhancement.

*2) Histogram Loss:* To ensure color consistency, we introduce a Histogram Loss to measure the discrepancy in color distribution between the enhanced image $I_{\text{gen}}$ and the corresponding reference image $I_{\text{gt}}$ in $R$, $G$, and $B$ channels, respectively. The Histogram Loss, $L_{\text{hist}}$, can be defined as

$$
\begin{aligned}
&L_{\text{hist}}(I_{\text{gen}}, I_{\text{gt}}) \\
&= \sum_{c=1}^{C} \left\| \frac{H(I_{\text{gen},c}, k)}{\sum_{k=1}^{K} H(I_{\text{gen},c}, k)} - \frac{H(I_{\text{gt},c}, k)}{\sum_{k=1}^{K} H(I_{\text{gt},c}, k)} \right\|_1 ,
\end{aligned} \tag{20}
$$

where the $H(I_{\text{gen},c}, k)$ and $H(I_{\text{gt},c}, k)$ are the histogram counts for the $k$-th bin in the $c$-th channel of the enhanced and reference images, respectively. $K$ denotes the total number of bins. The L1 norm $\|\cdot\|_1$ is used to calculate the sum of absolute differences between the normalized histograms of the enhanced and reference images across all bins and channels. This term ensures that the color distribution in the enhanced image closely matches that of the reference image, which is particularly important in challenging environments where color distortion is common.

*3) Combined Loss:* The overall loss function used to guide the training of our model combines the pixel-wise, edge, and histogram losses to ensure that the restored image maintains not only pixel-level accuracy but also structural integrity and accurate color representation.

The combined loss, $L$, is formulated as

$$
\begin{aligned}
L = &\lambda_1 L_{\text{pixel}}(I_{\text{gen}}, I_{\text{gt}}) + \lambda_2 L_{\text{edge}}(I_{\text{gen}}, I_{\text{gt}}) \\
&+ \lambda_3 L_{\text{hist}}(I_{\text{gen}}, I_{\text{gt}}).
\end{aligned} \tag{21}
$$

Here, $\lambda_1$, $\lambda_2$, and $\lambda_3$ are the learned weighting coefficients that dynamically balance the contributions of the Pixel-wise loss, Edge loss, and Histogram loss, respectively.

These weights allow the model to focus on different aspects of image quality–pixel-level accuracy, edge preservation, and color consistency–based on the specific degradation present in the input image.

## IV. EXPERIMENTS

To demonstrate the superior performance of our proposed ReviveDiff approach in enhancing image visibility under adverse conditions, we designed comprehensive experiments that benchmark our approach with SOTA approaches on five challenging tasks: image deraining, low-light image enhancement, nighttime dehazing, desmoking, and underwater image enhancement. Next, we first introduce the datasets and experiment details. Then, we present experimental comparisons with SOTA approaches and ablation studies on the key modules in our approach.

### A. Datasets

*1) Image Deraining:* For the image deraining task, we evaluated our ReviveDiff model using the Rain100L [38] dataset, which includes both clean and synthetic rainy images, with 200 pairs for training and 100 for testing.

*2) Underwater Image Enhancement:* For the underwater image enhancement task, we used three datasets: the UIEB dataset [59], the C60 [59] dataset, and the U45 dataset [59]. The UIEB dataset consists of 890 pairs of low-quality and high-quality images, while the C60 and U45 datasets comprise 60 and 45 challenging images, respectively, without reference images. Consistent with previous works, we split the UIEB dataset into 800 pairs for training and 90 images for testing (referred to as "T90"). We tested our ReviveDiff on all three datasets.

*3) Low-Light Image Enhancement:* For the low-light image enhancement task, we evaluated our ReviveDiff on the LOw-Light (LOL) [73] dataset, which comprises 500 image pairs of synthetic low-light and normal-light images. The LOL dataset provides 485 pairs for training and 15 pairs for testing.

*4) Nighttime Dehazing:* For the nighttime dehazing task, we used the GTA5 [5] dataset, a synthetic dataset created with the GTA5 game engine. The GTA5 dataset includes 864 image pairs, with 787 of them designated for training and the remaining 77 pairs for testing.

*5) Image Desmoking:* For the image desmoking task, we utilized the SMOKE [82] dataset, which consists of 132 paired real-world images captured in natural scenes with fog generated by a fog machine. 120 pairs of these images are used for training, and the remaining 12 pairs for testing.

### B. Implementation Details

During model training, we utilized a batch size of 4 and initialized the learning rate to $1 \times 10^{-4}$. We used the Adam optimizer [90], with parameters $\beta_1$ and $\beta_2$ to 0.9 and 0.999, for training optimization. The learning rate was adjusted using the MultiStepLR strategy for decay. The total number of training iterations varied based on the specific degradation being addressed. We employed a noise level of $\sigma = 90$ and set the diffusion step to 300 for all tasks. All experiments

TABLE I

COMPARISON OF IMAGE DERAINING RESULTS WITH SOTAS ON THE RAIN100L DERAINING DATASET [38]

| Methods | DerainNet [3] | SEMI [35] | UMRL [36] | DIDMDN [34] | Jorder [44] | MSPFN [45] | SPAIR [46] |
|---|---|---|---|---|---|---|---|
| PSNR ↑ | 27.03 | 25.03 | 29.18 | 25.23 | 36.61 | 32.40 | 37.30 |
| SSIM ↑ | 0.884 | 0.842 | 0.923 | 0.741 | 0.974 | 0.933 | 0.978 |
| LPIPS ↓ | — | — | — | — | 0.028 | — | — |
| Methods | MPRNet [37] | MAXIM [21] | AirNet [47] | IR-SDE [13] | MDARNet [15] | PromptIR [48] | M2PN [49] |
| PSNR ↑ | 36.40 | 38.06 | 34.90 | 38.30 | 35.68 | 37.04 | 38.36 |
| SSIM ↑ | 0.965 | 0.977 | 0.966 | 0.981 | 0.961 | 0.979 | 0.985 |
| LPIPS ↓ | 0.077 | 0.048 | — | 0.014 | — | — | — |
| Methods | NDR [50] | GANet [51] | FrePrompter [52] | AnyIR [53] | VLU-Net [54] | Perceive-IR [55] | **ReviveDiff (Ours)** |
| PSNR ↑ | 38.33 | 37.53 | 38.75 | 37.99 | 38.60 | 38.41 | 39.09 |
| SSIM ↑ | 0.984 | 0.973 | 0.984 | 0.982 | 0.984 | 0.984 | 0.979 |
| LPIPS ↓ | — | — | — | — | — | — | 0.012 |

TABLE II

QUANTITATIVE COMPARISON RESULTS OF UNDERWATER IMAGE ENHANCEMENT WITH SOTAS ON THE T90 [59], U45 [59], AND C60 [59] UNDERWATER DATASETS

| Dataset | Methods | GDCP [60] | UGAN [61] | FUnIEGAN [62] | UWCNN [56] | Ucolor [63] | MLLE [18] | UShape [8] | FA+Net [9] | NU2Net [10] | **ReviveDiff** |
|---|---|---|---|---|---|---|---|---|---|---|---|
| T90 | PSNR ↑ | 13.89 | 17.42 | 16.97 | 19.02 | 20.86 | 19.48 | 20.24 | 20.98 | 22.93 | 25.01 |
| | SSIM ↑ | 0.75 | 0.76 | 0.73 | 0.82 | 0.88 | 0.84 | 0.81 | 0.88 | 0.90 | 0.92 |
| | NIQE ↓ | 4.93 | 5.81 | 4.92 | 4.72 | 4.75 | 4.87 | 4.67 | 4.83 | 4.81 | 4.45 |
| U45 | UCIQE ↑ | 0.59 | 0.57 | 0.56 | 0.48 | 0.58 | 0.59 | 0.57 | 0.58 | 0.57 | 0.62 |
| | NIQE ↓ | 4.14 | 5.79 | 4.45 | 4.09 | 4.71 | 4.83 | 4.20 | 4.02 | 5.7 | 3.93 |
| C60 | UCIQE ↑ | 0.57 | 0.55 | 0.55 | 0.51 | 0.55 | 0.57 | 0.56 | 0.57 | 0.58 | 0.59 |
| | NIQE ↓ | 6.27 | 6.90 | 6.06 | 5.94 | 6.14 | 5.85 | 5.60 | 5.70 | 5.64 | 5.57 |

were conducted on a single NVIDIA RTX 4090 GPU using the PyTorch platform.

### C. Evaluation Metrics

In our experiments, we adopt several widely used evaluation metrics to assess our model's performance across various restoration tasks comprehensively. We employ the Peak Signal-to-Noise Ratio (PSNR) [63] and the Structural Similarity Index Measure (SSIM) [64] to quantify the fidelity of the restored images compared to the ground truth. Additionally, the Learned Perceptual Image Patch Similarity (LPIPS) [91] metric is used to evaluate perceptual similarity, reflecting human visual perception. For tasks such as underwater image enhancement, where reference images may not be available, we further utilize no-reference metrics like the Natural Image Quality Evaluator (NIQE) [92] and the Underwater Color Image Quality Evaluation (UCIQE) [93] to assess naturalness and color accuracy. Together, these metrics provide a robust framework for evaluating both pixel-level accuracy and perceptual quality across diverse adverse conditions.

### D. Comparisons With the State of the Arts

*1) Image Deraining:* Images captured in rainy conditions often suffer from the detrimental effects of rain streaks. To evaluate the deraining capability of our method, we utilized three widely used metrics: PSNR, SSIM, and LPIPS. We compared our method with several SOTAs from 2016 to 2025.

As shown in Table I, our ReviveDiff model achieves a competitive performance across all these metrics. When compared to the second-best method, FrePrompter [52], our PSNR surpassed it by 0.34, making ReviveDiff the only method to achieve a PSNR over 39. This indicates the superior quantitative performance of our ReviveDiff in deraining. Visually, as depicted in Fig. 5, our results show a significant reduction in rain presence and are most closely aligned with the reference image in terms of quality.

*2) Underwater Image Enhancement:* Underwater images often suffer from degradation due to the scattering and absorption effects of water, coupled with low illumination, resulting in poor visual quality. In our study, we evaluated ReviveDiff using both paired (T90) and non-reference datasets (U45 and C60). For the T90 dataset, we employed widely recognized metrics: PSNR and SSIM, to assess image quality. Additionally, we utilized two metrics that do not require reference images for computation: UCIQE (to evaluate brightness) and NIQE (to assess distortion and noise). These metrics were also applied to the U45 and C60 datasets. We compared our ReviveDiff model with SOTAs [57], among which GDCP [60] represents traditional theory-based approaches, while the others are deep learning-based.

| T90: Input | UWCNN [56] | FA+Net [9] | NU2Net [10] | **ReviveDiff (Ours)** | GT |

(a) With-Reference Results



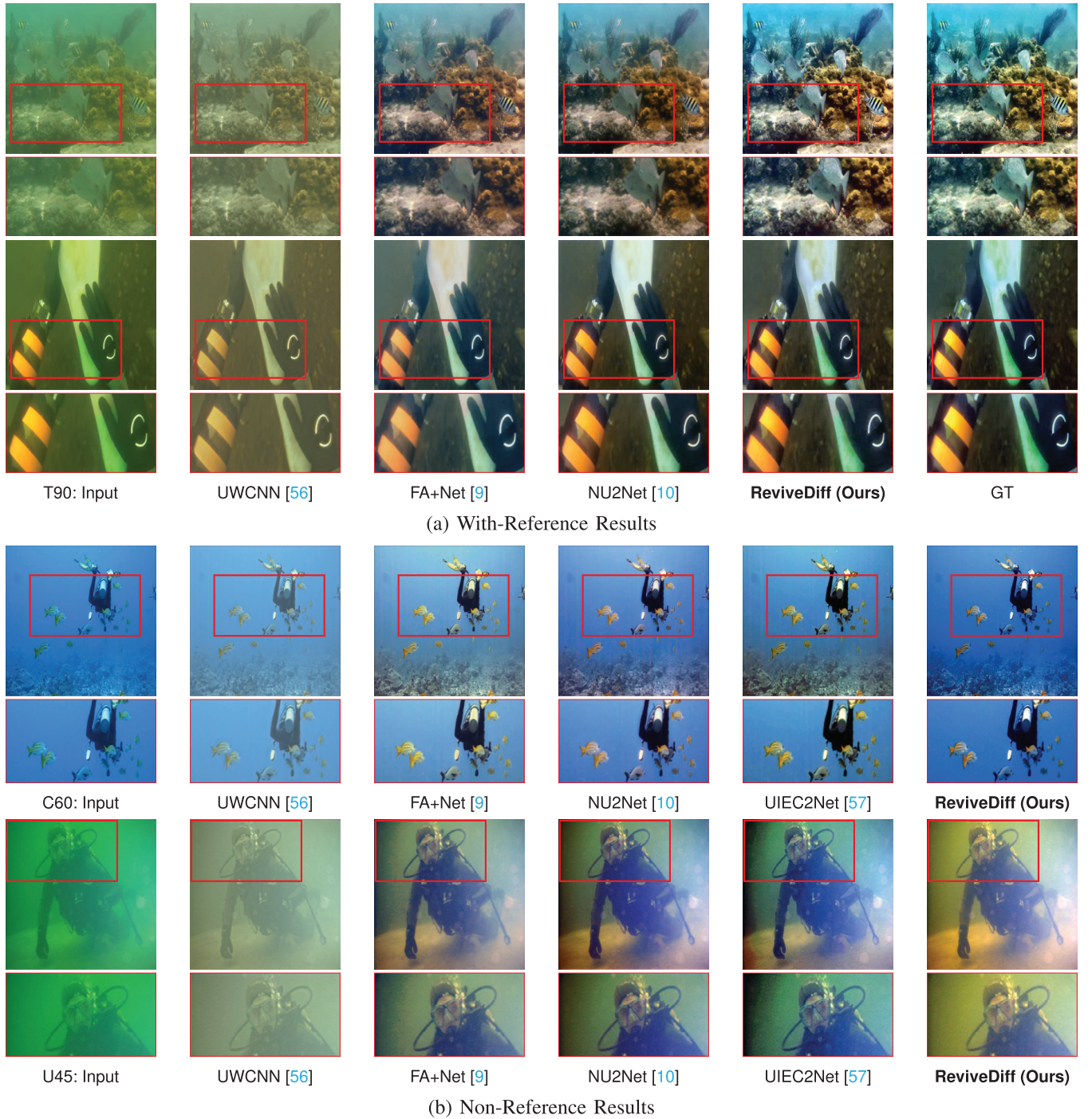| C60: Input | UWCNN [56] | FA+Net [9] | NU2Net [10] | UIEC2Net [57] | **ReviveDiff (Ours)** |
| U45: Input | UWCNN [56] | FA+Net [9] | NU2Net [10] | UIEC2Net [57] | **ReviveDiff (Ours)** |

(b) Non-Reference Results

Fig. 6. Visual comparison of underwater image enhancement results with SOTAs on T90 [59], U45 [59], and C60 [59].

For a quantitative comparison, as shown in Table II, our ReviveDiff outperforms all other methods across all three datasets. Specifically, ReviveDiff achieves exceptional performance in PSNR and SSIM metrics. Compared to the second-best method, NU2Net [10], our PSNR is higher by 2.08, and ReviveDiff is the only method to achieve an SSIM greater than 0.9. Furthermore, on the U45 dataset, the proposed ReviveDiff is the only method to achieve an NIQE score lower than 4, indicating its superior restoration performance of brightness.

For visual quality comparison, as illustrated in Fig. 6, our ReviveDiff generates results that are most similar to the reference images in the T90 dataset. Additionally, for the non-reference datasets U45 and C60, our results also demonstrate better visual quality.

*3) Low-Light Image Enhancement:* Enhancing the image quality of low-light images is a significant challenge in current research. To assess the performance of our ReviveDiff model, we employed metrics such as PSNR, SSIM, and LPIPS and compared our results with SOTA methods from 2018 to 2024.
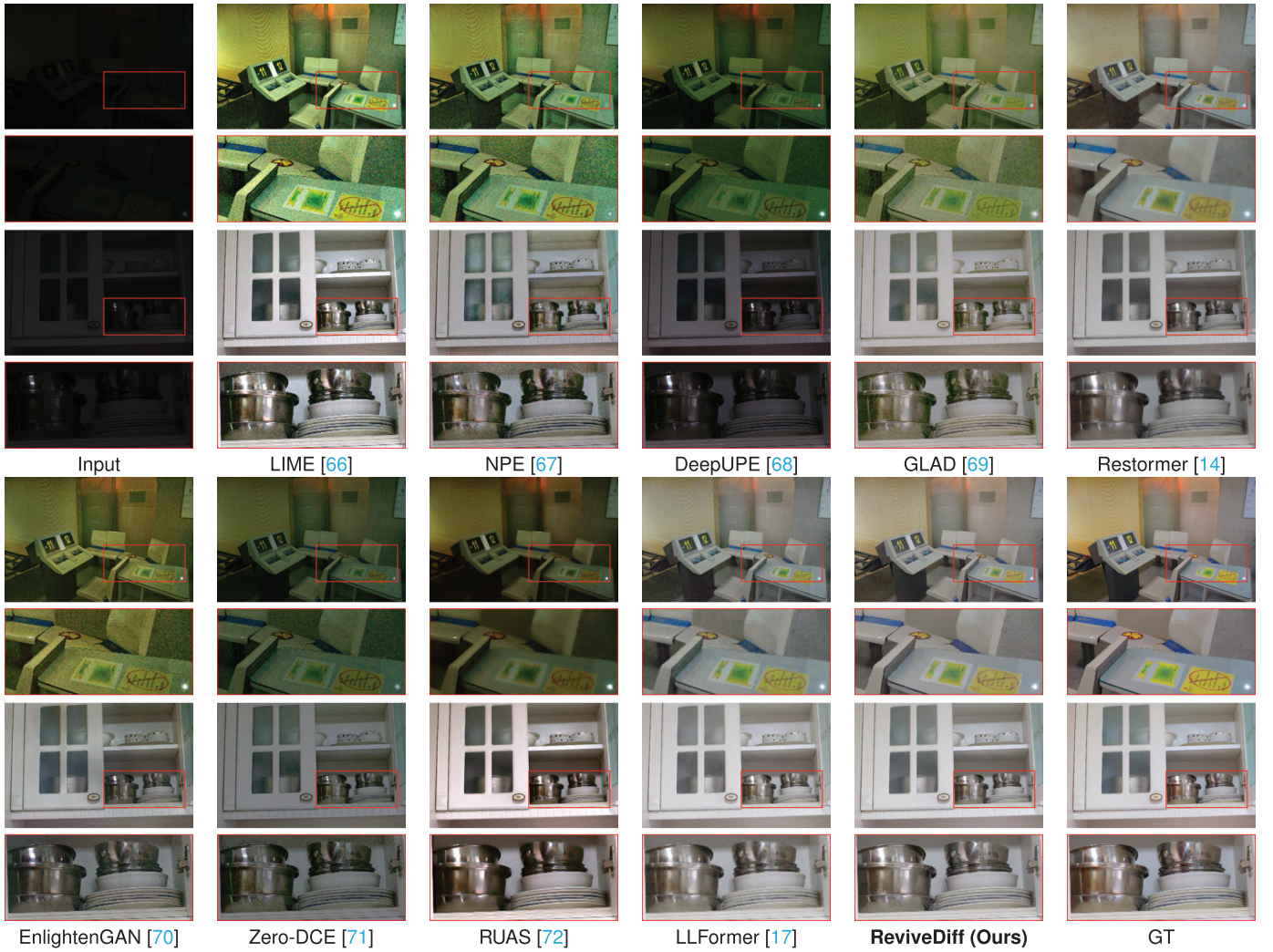
Fig. 7. Visual comparison of low-light image enhancement results with SOTAs on the LOL dataset [73]. The small images enclosed in red below each result provide an enlarged view of the areas highlighted by red rectangles.

TABLE III
QUANTITATIVE COMPARISONS OF LOW-LIGHT IMAGE ENHANCEMENT RESULTS ON THE LOL DATASET [73]

| Methods | RetinexNet [2] | KIND [74] | DSLR [75] | DRBN [76] | Zero-DCE [71] | MIRNet [77] | EnlightenGAN [70] |
|---|---|---|---|---|---|---|---|
| PSNR ↑ | 16.774 | 20.84 | 14.816 | 16.774 | 14.816 | 24.138 | 17.606 |
| SSIM ↑ | 0.462 | 0.790 | 0.572 | 0.462 | 0.572 | 0.830 | 0.653 |
| LPIPS ↓ | 0.417 | 0.170 | 0.375 | 0.417 | 0.375 | 0.250 | 0.372 |

| Methods | RUAS [72] | KIND++ [78] | DDIM [27] | CDEF [79] | SCI [80] | URetinex-Net [2] | Uformer [20] |
|---|---|---|---|---|---|---|---|
| PSNR ↑ | 16.405 | 21.300 | 16.521 | 16.335 | 14.784 | 19.842 | 19.001 |
| SSIM ↑ | 0.503 | 0.820 | 0.776 | 0.585 | 0.525 | 0.824 | 0.741 |
| LPIPS ↓ | 0.364 | 0.160 | 0.376 | 0.407 | 0.366 | 0.237 | 0.354 |

| Methods | Restormer [14] | UHDFour [6] | WeatherDiff [81] | LLformer [17] | LightenDiffusion [29] | PromptGIP [22] | **ReviveDiff (Ours)** |
|---|---|---|---|---|---|---|---|
| PSNR ↑ | 20.614 | 23.093 | 17.913 | 23.650 | 20.453 | 20.30 | 24.272 |
| SSIM ↑ | 0.797 | 0.821 | 0.811 | 0.816 | 0.803 | 0.803 | 0.832 |
| LPIPS ↓ | 0.288 | 0.259 | 0.272 | 0.171 | 0.192 | N/A | 0.0875 |

As shown in Table III, our ReviveDiff model outperforms competing approaches across these metrics. Notably, our method is the only one to achieve an LPIPS score lower than 0.1, signifying unparalleled results in terms of human visual perception. Furthermore, as depicted in Fig. 7, our results demonstrate superior visual quality, exhibiting minimal color distortion and optimal brightness, closely matching the ground truth images.
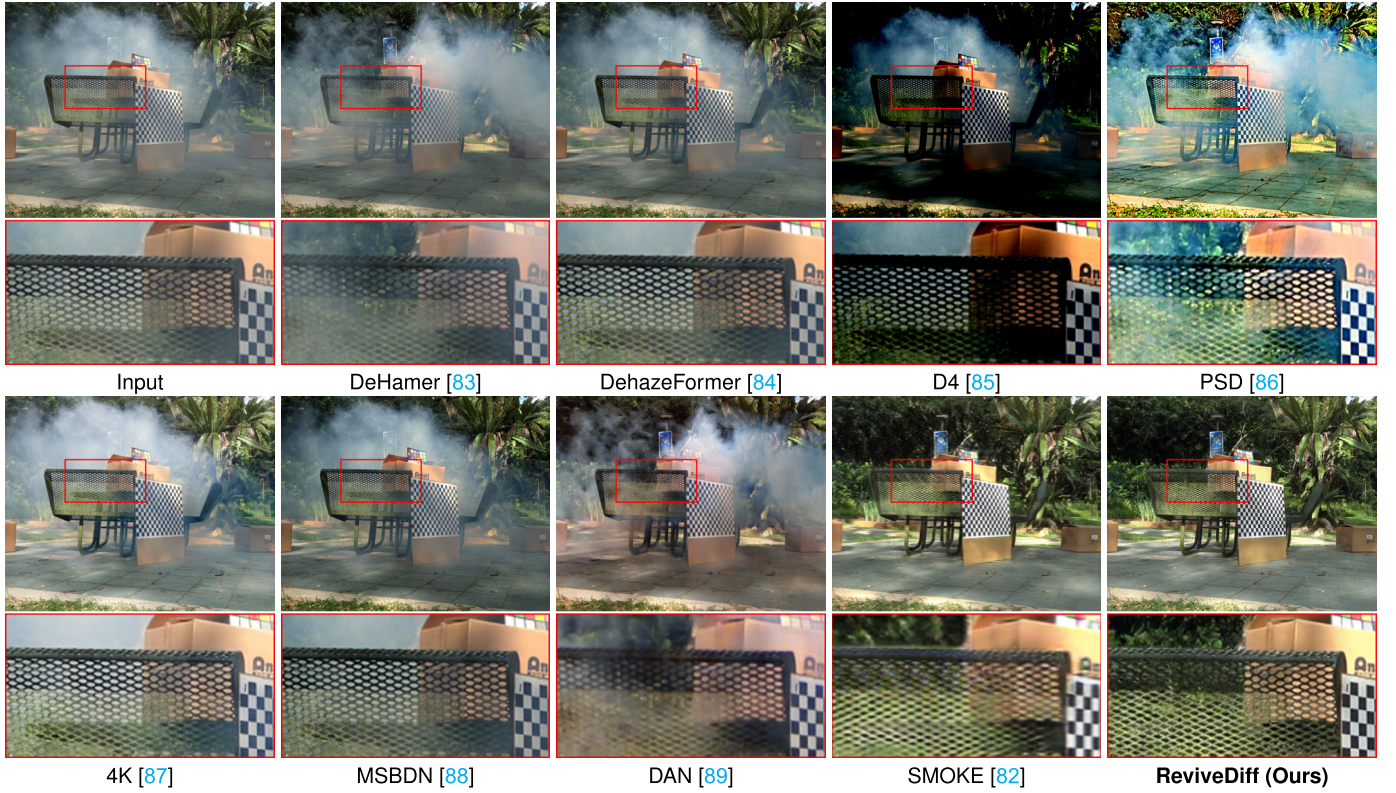
Fig. 8. Visual comparison of smoke image enhancement results with SOTAs on the SMOKE dataset [82].

TABLE IV
COMPARISON OF IMAGE DESMOKING RESULTS ON THE SMOKE DATASET [82]

| Methods | PSNR ↑ | SSIM ↑ |
|---|---|---|
| DCP [94] | 11.26 | 0.26 |
| GDN [95] | 15.19 | 0.53 |
| MSBDN [88] | 13.19 | 0.34 |
| DeHamer [83] | 13.31 | 0.28 |
| SMOKE [82] | 18.83 | 0.62 |
| CTHD-Net [96] | 19.23 | 0.63 |
| **ReviveDiff (Ours)** | 20.09 | 0.65 |

TABLE V
ABLATION STUDIES ON THE UIEB UNDERWATER DATASET [59]

| Methods | PSNR ↑ | SSIM ↑ | NIQE ↓ |
|---|---|---|---|
| Only $L_1$ Loss | 24.47 | 0.91 | 4.61 |
| $w/o$ Edge Prior | 24.26 | 0.92 | 4.51 |
| $w/o$ His Prior | 24.25 | 0.92 | 4.48 |
| $w/o$ MAFC | 20.08 | 0.86 | 4.65 |
| $w/o$ C2F | 24.04 | 0.92 | 4.51 |
| $W_1 : W_2 = 1:1$ | 24.64 | 0.92 | 4.58 |
| $W_1 : W_2 = 10:1$ | 24.55 | 0.92 | 4.56 |
| $W_1 : W_2 = 1:10$ | 23.58 | 0.91 | 4.53 |
| **ReviveDiff (Ours)** | 25.01 | 0.92 | 4.45 |

TABLE VI
ABLATION STUDIES ON THE GTA5 NIGHTTIME DEHAZING DATASET [5]

| Methods | PSNR ↑ | SSIM ↑ | LPIPS ↓ |
|---|---|---|---|
| Zhang *et al.* [97] | 20.92 | 0.646 | - |
| Ancuti *et al.* [98] | 20.59 | 0.623 | - |
| Yan *et al.* [7] | 27.00 | 0.850 | - |
| CycleGAN [99] | 21.75 | 0.696 | - |
| Jin *et al.* [5] | 30.38 | 0.904 | 0.099 |
| Only $L_1$ Loss | 30.82 | 0.840 | 0.103 |
| $w/o$ Edge Prior | 30.56 | 0.838 | 0.107 |
| $w/o$ His Prior | 30.49 | 0.837 | 0.106 |
| $w/o$ MAFC | 23.77 | 0.722 | 0.206 |
| $w/o$ C2F | 32.52 | 0.840 | 0.098 |
| **ReviveDiff (Ours)** | 32.91 | 0.895 | 0.094 |

preserve the low-light environment while simultaneously removing haze. As demonstrated in Table VI, our ReviveDiff model surpasses SOTAs in both PSNR and LPIPS metrics, experiencing a marginal loss in SSIM by only 0.009 (1%) compared to the method by Jin et al. [5]. However, our ReviveDiff model outperforms that of Jin et al. in PSNR by 2.53 (8%), and a lower LPIPS score indicates better alignment with human perception. This is further supported by visual quality comparisons in Fig. 9, where it is evident that the method by Jin et al. [5] fails to restore light as effectively as ours, resulting in darker images with lower contrast.

*4) Nighttime Image Dehazing:* Nighttime image Dehazing presents a highly complex task, requiring methods that

*5) Real-World Image Desmoking:* Real-world image desmoking is a challenging task that requires methods

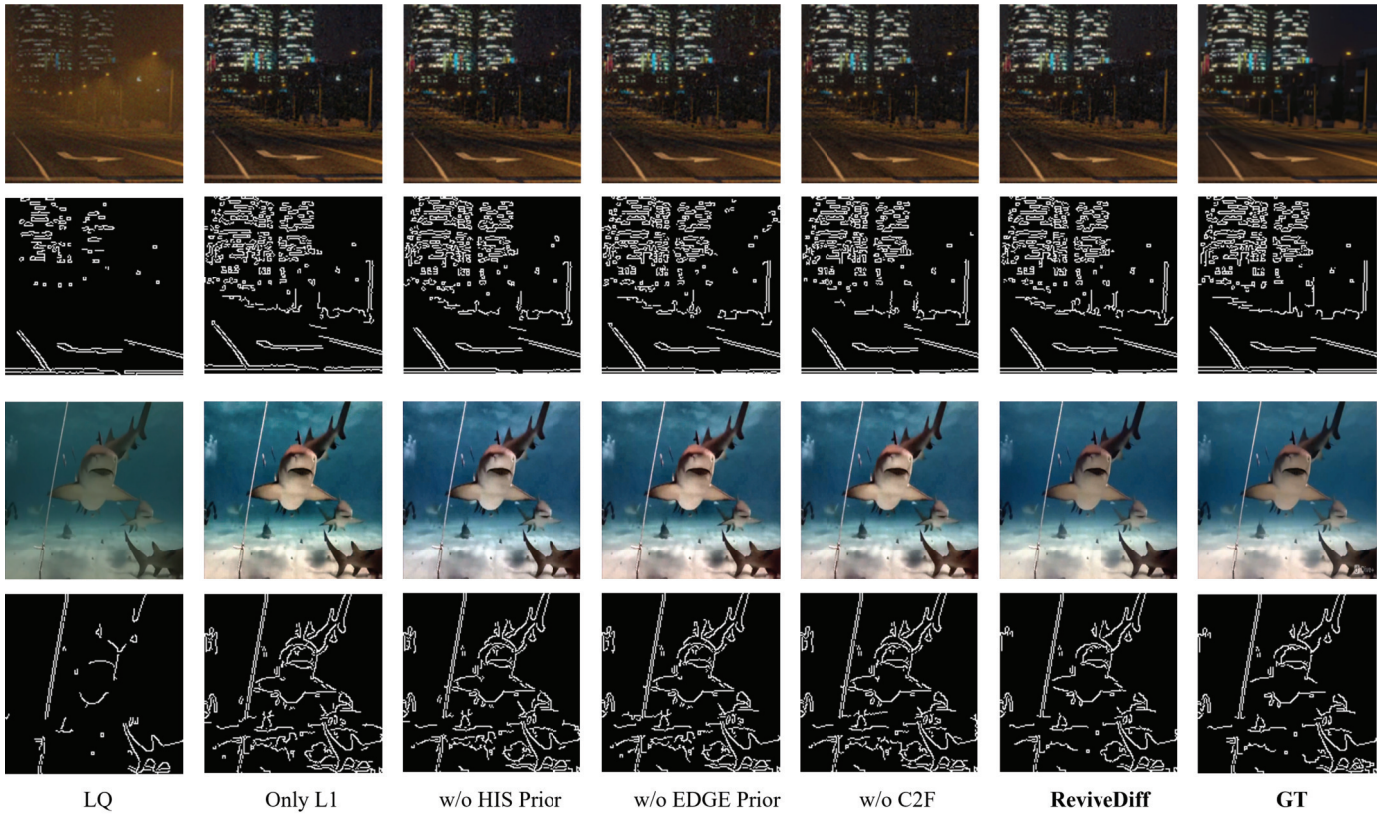| LQ | Only L1 | w/o HIS Prior | w/o EDGE Prior | w/o C2F | **ReviveDiff** | **GT** |

Fig. 9. Visual comparison of the ablation study results and edge maps on the GTA5 nighttime hazing dataset [5] and the UIEB underwater image dataset [59].

capable of effectively removing dense fog while preserving the underlying image details and maintaining natural visibility.

Table IV shows a quantitative comparison of desmoking performance. As shown in this table, our method achieves the highest PSNR of 20.09 dB and SSIM of 0.65, outperforming other SOTA methods such as SMOKE [82] (PSNR 18.83 dB, SSIM 0.62) and CTHD-Net [96] (PSNR 19.23 dB, SSIM 0.63). This significant improvement highlights the superior capability of our method in restoring clear images while preserving structural information. The results demonstrate that our approach surpasses existing techniques, achieving state-of-the-art performance in the image-desmoking task.

### E. Ablation Studies

We performed additional ablation studies to validate the effectiveness of our key components: Coarse-to-Fine Learning (C2F), Multi-Attentional Feature Complementation (MAFC), and the guided losses based on Edge and Histogram Priors. Quantitative comparisons on the UIEB underwater dataset and the GTA5 nighttime hazing dataset are presented in Tables V and VI, respectively. Fig. 9 provides visual comparisons, including edge maps, highlighting the influence of each ablated component in these challenging scenarios.

*1) Effectiveness of Coarse-to-Fine Learning:* Tables V and VI show that even without Coarse-to-Fine Learning (denoted as *w/o C2F*), our model still outperforms some competing methods. However, it falls short of the full ReviveDiff model in terms of PSNR, SSIM, LPIPS, and NIQE, confirming

the importance of progressively capturing both coarse and fine features for robust enhancement in nighttime dehazing and underwater tasks. Moreover, as the edge maps illustrate, omitting Coarse-to-Fine Learning leads to significant detail loss in nighttime dehazing images in underwater images. For instance, in the visualization of the underwater image, the result of *w/o C2F* exhibits inconsistent colors. Moreover, its edge map of the nighttime defogging image loses some structural details; conversely, the edge map of the underwater image reinforces unnecessary details.

*2) Effectiveness of Multi-Attentional Feature Complementation:* The variant without MAFC (denoted as *w/o MAFC*), which replaces MAFC with a simple addition operation, suffers a significant performance drop in both Table V and Table VI. For instance, on the UIEB dataset (Table V), removing MAFC reduces PSNR from 25.01 dB to 20.08 dB and SSIM from 0.92 to 0.86, indicating poorer overall restoration. On the GTA5 dataset (Table VI), it similarly lowers performance by around 0.4 dB in PSNR and increases LPIPS by over 0.01 compared to the full ReviveDiff model. These declines highlight MAFC's critical role in balancing and fusing coarse and fine features effectively.

*3) Discussion on Dynamic Weighting in MAFC:* To further validate the effectiveness of the proposed dynamic weighting strategy within the MAFC module, we conducted an additional ablation study on the UIEB dataset by comparing three fixed weight ratios ($W_1 : W_2$) against our adaptive weighting approach. Specifically, we tested: (1) $W_1 : W_2 = 1 : 1$,

TABLE VII

EFFICIENCY COMPARISONS BETWEEN SIMILAR SOTAS AND OUR REVIVEDIFF, THE PARAMETERS AND MULTI-ADDS ARE COMPUTED WITH AN INPUT SIZE OF $3 \times 256 \times 256$

| Dataset | Methods | Parameters | Multi-Adds | PSNR/SSIM/LPIPs(NIQE) |
|---|---|---|---|---|
| Rain100L [38] | IR-SDE [13] | 137.15M | 379.31G | 38.30 / **0.981** / 0.014 |
| | AWRaCLe [100] | 94.25M | 158.64G | 35.71 / 0.966 / - |
| | **ReviveDiff (Ours)** | **88.31M** | **73.95G** | **39.09** / <u>0.979</u> / **0.012** |
| Nighttime [5] | IR-SDE [13] | 137.15M | 379.31G | 31.34 / 0.856 / 0.098 |
| | **ReviveDiff (Ours)** | **88.31M** | **73.95G** | **32.91 / 0.895 / 0.094** |
| UIEB [58] | IR-SDE [13] | 137.15M | 379.31G | 21.55 / 0.89 / 4.75 |
| | **ReviveDiff (Ours)** | **88.31M** | **73.95G** | **25.01 / 0.92 / 4.45** |
| LOL [73] | WeatherDiff [81] | 87.90M | 261.55G | 17.913 / 0.811 / 0.272 |
| | **ReviveDiff (Ours)** | **88.31M** | **73.95G** | **24.272 / 0.832 / 0.0875** |

(2) $W_1 : W_2 = 10 : 1$, and (3) $W_1 : W_2 = 1 : 10$. As shown in Table V, all fixed-weight variants exhibit noticeable performance degradation compared to our dynamic weighting strategy. For instance, the equal-weight assignment $(1 : 1)$ achieves a PSNR of 24.64 dB and a NIQE of 4.58, whereas the highly skewed ratio $(1 : 10)$ further degrades the PSNR to 23.58 dB and SSIM to 0.91. In contrast, our adaptive weighting mechanism reaches a superior PSNR of 25.01 dB and a lower NIQE of 4.45. These quantitative results explicitly validate that dynamically adjusting the weighting factors between coarse and fine features significantly enhances performance, confirming the critical role of adaptive weighting in addressing diverse degradations present in underwater images.

*4) Effectiveness of the Prior-Guided Losses:* Fig. 9 and Tables V and VI present the results for *w/o Edge Prior* and *w/o His Prior*, demonstrating that omitting these priors degrades both boundary preservation and color fidelity. In particular, the *Only $L_1$ Loss* variant shows that relying solely on a pixel-based metric is insufficient for handling complex degradations, often yielding unnatural color distributions and less-defined edges. As seen in Fig. 9, *w/o Edge Prior* fails to maintain sharp object boundaries. At the same time, *w/o His Prior* leads to noticeable color inconsistencies–especially in underwater images, where the restoration process suffers from an uncorrected color cast. In contrast, incorporating both Edge and Histogram priors simultaneously sharpens edges and refines color accuracy. The improved edge maps in the lower rows of Fig. 9 verify that these guided losses help preserve subtle details and rectify color distortions under challenging low-visibility conditions. This highlights how the proposed prior-guided losses address the limitations of purely pixel-based objectives, ultimately enabling more robust restoration in adverse environments.

*5) Efficiency Analysis:* Table VII provides a comprehensive efficiency comparison of our proposed ReviveDiff model against similar state-of-the-art methods, including IR-SDE [13], WeatherDiff [81], and AWRaCLe [100], in terms of model complexity (number of parameters), computational cost (Multi-Adds), and restoration performance across multiple datasets. Our proposed ReviveDiff significantly reduces computational overhead, requiring only 88.31M parameters and 73.95G Multi-Adds. Despite this reduced complexity, ReviveDiff consistently achieves superior performance compared to other methods, notably surpassing IR-SDE on Rain100L (39.09 dB vs. 38.30 dB PSNR), Nighttime dehazing

(32.91 dB vs. 31.34 dB PSNR), and UIEB underwater datasets (25.01 dB vs. 21.55 dB PSNR). Additionally, ReviveDiff significantly outperforms WeatherDiff on the LOL dataset (24.272 dB vs. 17.913 dB PSNR) and AWRaCLe on Rain100L (39.09 dB vs. 35.71 dB PSNR). These results underscore ReviveDiff's efficiency and effectiveness, demonstrating its suitability for practical deployment in resource-constrained environments.

## V. CONCLUSION

In this paper, we developed a novel diffusion model, "ReviveDiff", specifically designed for enhancing images degraded by diverse adverse environmental conditions such as fog, rain, underwater scenarios, and nighttime settings. ReviveDiff features a unique Coarse-to-Fine Learning framework and a Multi-Attention Feature Complementation module, enabling the effective learning and fusion of multi-scale features. Additionally, it incorporates edge and histogram priors for enhanced structure and color restoration. Upon evaluation across seven datasets covering a range of adverse degradations, ReviveDiff demonstrably outperforms existing SOTA methods, offering a robust solution for universal adverse condition image enhancement challenges. To our knowledge, ReviveDiff represents a breakthrough contribution by utilizing a diffusion-based universal approach to address adverse conditions. This work not only underscores the versatility and efficacy of diffusion models in restoring images affected by adverse conditions but also establishes a new benchmark for future research. Looking ahead, we plan to explore strategies for reducing the number of diffusion steps and accelerating inference, further enhancing the model's efficiency without compromising its restoration quality.

## REFERENCES

[1] J. Liu, Q. Wang, H. Fan, Y. Wang, Y. Tang, and L. Qu, "Residual denoising diffusion models," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, vol. 32, Jun. 2024, pp. 2773–2783.

[2] W. Chen, W. Wang, W. Yang, and J. Liu, "Deep retinex decomposition for low-light enhancement," in *Proc. Brit. Mach. Vis. Conf.*, Jan. 2018, pp. 1–11.

[3] X. Fu, J. Huang, X. Ding, Y. Liao, and J. Paisley, "Clearing the skies: A deep network architecture for single-image rain removal," *IEEE Trans. Image Process.*, vol. 26, no. 6, pp. 2944–2956, Jun. 2017.

[4] E. H. Land, "The retinex theory of color vision," *Scientific Amer.*, vol. 237, no. 6, pp. 108–129, Dec. 1977.

[5] Y. Jin, B. Lin, W. Yan, Y. Yuan, W. Ye, and R. T. Tan, "Enhancing visibility in nighttime haze images using guided APSF and gradient adaptive convolution," in *Proc. 31st Int. Conf. Multimedia*, Oct. 2023, pp. 2446–2457.

[6] C. Li et al., "Embedding Fourier for Ultra-High-Definition low-light image enhancement," 2023, *arXiv:2302.11831*.

[7] W. Yan, R. T. Tan, and D. Dai, "Nighttime defogging using high-low frequency decomposition and grayscale-color networks," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, 2020, pp. 473–488.

[8] L. Peng, C. Zhu, and L. Bian, "U-shape transformer for underwater image enhancement," *IEEE Trans. Image Process.*, vol. 32, pp. 3066–3079, 2023.

[9] J. Jiang et al., "Five a+ network: You only need 9k parameters for underwater image enhancement," 2023, *arXiv:2305.08824*.

[10] C. Guo et al., "Underwater ranker: Learn which is better and how to be better," in *Proc. AAAI Conf. Artif. Intell.*, 2023, pp. 702–709.

[11] D. Du, E. Li, L. Si, F. Xu, J. Niu, and F. Sun, "UIEDP: Underwater image enhancement with diffusion prior," 2023, *arXiv:2312.06240*.

[12] L. Chen, X. Chu, X. Zhang, and J. Sun, "Simple baselines for image restoration," in *Proc. Eur. Conf. Comput. Vis.*, Jan. 2022, pp. 17–33.

[13] Z. Luo, F. Gustafsson, Z. Zhao, J. Sjölund, and T. B. Schön, "Image restoration with mean-reverting stochastic differential equations," in *Proc. Int. Conf. Mach. Learn.*, Jan. 2023, pp. 23045–23066.

[14] S. W. Zamir, A. Arora, S. Khan, M. Hayat, F. S. Khan, and M.-H. Yang, "Restormer: Efficient transformer for high-resolution image restoration," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2022, pp. 5728–5739.

[15] Z. Hao, S. Gai, and P. Li, "Multi-scale self-calibrated dual-attention lightweight residual dense deraining network based on monogenic wavelets," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 33, no. 6, pp. 2642–2655, Jun. 2023.

[16] D. Ren, W. Zuo, Q. Hu, P. Zhu, and D. Meng, "Progressive image deraining networks: A better and simpler baseline," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 3932–3941.

[17] T. Wang, K. Zhang, T. Shen, W. Luo, B. Stenger, and T. Lu, "Ultra-high-definition lowlight image enhancement: A benchmark and transformer-based method," in *Proc. AAAI Conf. Artif. Intell.*, vol. 37, no. 3, 2023, pp. 2654–2662.

[18] W. Zhang, P. Zhuang, H.-H. Sun, G. Li, S. Kwong, and C. Li, "Underwater image enhancement via minimal color loss and locally adaptive contrast enhancement," *IEEE Trans. Image Process.*, vol. 31, pp. 3997–4010, 2022.

[19] Y. Cui et al., "Selective frequency network for image restoration," in *Proc. Int. Conf. Learn. Represent.*, 2023, pp. 1–11.

[20] Z. Wang, X. Cun, J. Bao, W. Zhou, J. Liu, and H. Li, "Uformer: A general U-Shaped transformer for image restoration," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2022, pp. 17662–17672.

[21] Z. Tu et al., "MAXIM: Multi-axis MLP for image processing," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2022, pp. 5769–5780.

[22] Y. Liu et al., "Unifying image processing as visual prompting question answering," in *Proc. Int. Conf. Mach. Learn.*, Jan. 2023, pp. 30873–30891.

[23] D. Zheng, X.-M. Wu, S. Yang, J. Zhang, J.-F. Hu, and W. Zheng, "Selective hourglass mapping for universal image restoration based on diffusion model," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2024, pp. 25445–25455.

[24] Z. Luo, F. K. Gustafsson, Z. Zhao, J. Sjölund, and T. B. Schön, "Controlling vision-language models for multi-task image restoration," in *Proc. Int. Conf. Learn. Represent.*, 2024, pp. 1–21.

[25] Y. Liu, Z. Ke, F. Liu, N. Zhao, and R. W. H. Lau, "Diff-plugin: Revitalizing details for diffusion-based low-level tasks," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2024, pp. 4197–4208.

[26] S. Sun, W. Ren, X. Gao, R. K. Wang, and X. Cao, "Restoring images in adverse weather conditions via histogram transformer," in *Proc. Eur. Conf. Comput. Vis.*, Sep. 2024, pp. 111–129.

[27] J. Song, C. Meng, and S. Ermon, "Denoising diffusion implicit models," 2020, *arXiv:2010.02502*.

[28] B. Fei et al., "Generative diffusion prior for unified image restoration and enhancement," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2023, pp. 9935–9946.

[29] H. Jiang, A. Luo, X. Liu, S. Han, and S. Liu, "LightenDiffusion: Unsupervised low-light image enhancement with latent-retinex diffusion models," in *Proc. Eur. Conf. Comput. Vis.*, Nov. 2024, pp. 161–179.

[30] Y. Tang, H. Kawasaki, and T. Iwaguchi, "Underwater image enhancement by transformer-based diffusion model with non-uniform sampling for skip strategy," in *Proc. 31st ACM Int. Conf. Multimedia*, Oct. 2023, pp. 5419–5427.

[31] C. Wang et al., "Reconciling stochastic and deterministic strategies for zero-shot image restoration using diffusion model in dual," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2025, pp. 23207–23216.

[32] S. Kim, S. Suh, and M. Lee, "RAD: Region-aware diffusion models for image inpainting," 2024, *arXiv:2412.09191*.

[33] Z. Luo, F. Gustafsson, Z. Zhao, J. Sjölund, and T. B. Schön, "Refusion: Enabling large-size realistic image restoration with latent-space diffusion models," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops*, Jun. 2023, pp. 1680–1691.

[34] H. Zhang and V. M. Patel, "Density-aware single image de-raining using a multi-stream dense network," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 695–704.

[35] W. Wei, D. Meng, Q. Zhao, Z. Xu, and Y. Wu, "Semi-supervised transfer learning for image rain removal," 2018, *arXiv:1807.11078*.

[36] R. Yasarla and V. M. Patel, "Uncertainty guided multi-scale residual learning-using a cycle spinning CNN for single image de-raining," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 8397–8406.

[37] S. W. Zamir et al., "Multi-stage progressive image restoration," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2021, pp. 14821–14831.

[38] W. Yang, R. T. Tan, J. Feng, J. Liu, Z. Guo, and S. Yan, "Deep joint rain detection and removal from a single image," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 1685–1694.

[39] Y. Dai, F. Gieseke, S. Oehmcke, Y. Wu, and K. Barnard, "Attentional feature fusion," in *Proc. IEEE Winter Conf. Appl. Comput. Vis. (WACV)*, Jan. 2021, pp. 3560–3569.

[40] Z. Chen, Z. He, and Z.-M. Lu, "DEA-net: Single image dehazing based on detail-enhanced convolution and content-guided attention," *IEEE Trans. Image Process.*, vol. 33, pp. 1002–1015, 2024.

[41] S. Woo, J. Park, J. Lee, and I. S. Kweon, "CBAM: Convolutional block attention module," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, Jan. 2018, pp. 3–19.

[42] Q. Wang, B. Wu, P. Zhu, P. Li, W. Zuo, and Q. Hu, "ECA-Net: Efficient channel attention for deep convolutional neural networks," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2020, pp. 11534–11542.

[43] H. Zhao, X. Kong, J. He, Y. Qiao, and C. Dong, "Efficient image super-resolution using pixel attention," in *Proc. Eur. Conf. Comput. Vis.*, Oct. 2020, pp. 56–72.

[44] W. Yang, R. T. Tan, J. Feng, Z. Guo, S. Yan, and J. Liu, "Joint rain detection and removal from a single image with contextualized deep networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 42, no. 6, pp. 1377–1393, Jun. 2020.

[45] Y. Zhang, Q. Li, M. Qi, D. Liu, J. Kong, and J. Wang, "Multi-scale frequency separation network for image deblurring," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 33, no. 10, pp. 5525–5537, Oct. 2023.

[46] K. Purohit, M. Suin, A. N. Rajagopalan, and V. Naresh Boddeti, "Spatially-adaptive image restoration using distortion-guided networks," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2021, pp. 2289–2299.

[47] B. Li, X. Liu, P. Hu, Z. Wu, J. Lv, and X. Peng, "All-In-One image restoration for unknown corruption," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2022, pp. 17431–17441.

[48] V. Potlapalli, S. W. Zamir, S. Khan, and F. S. Khan, "PromptIR: Prompting for All-in-One blind image restoration," in *Proc. Int. Conf. Neural Inf. Process. Syst.*, Jan. 2023, pp. 71275–71293.

[49] J. Zhong, Z. Li, D. Xiang, M. Han, C. Li, and Y. Gan, "A lightweight multi-domain multi-attention progressive network for single image deraining," in *Proc. ACM Int. Conf. Multimedia*, Oct. 2024, pp. 2001–2010.

[50] M. Yao, R. Xu, Y. Guan, J. Huang, and Z. Xiong, "Neural degradation representation learning for All-in-One image restoration," *IEEE Trans. Image Process.*, vol. 33, pp. 5408–5423, 2024.

[51] Y. Yuan, X. Guo, and D. Ma, "Global attention network with rain prior for real time single image deraining," *Neurocomputing*, vol. 625, Apr. 2025, Art. no. 129490.

[52] Z. Wu, W. Liu, J. Wang, J. Li, and D. Huang, "FrePrompter: Frequency self-prompt for all-in-one image restoration," *Pattern Recognit.*, vol. 161, May 2025, Art. no. 111223.

[53] B. Ren et al., "Any image restoration via efficient spatial-frequency degradation adaptation," 2025, *arXiv:2504.14249*.

[54] H. Zeng, X. Wang, Y. Chen, J. Su, and J. Liu, "Vision-language gradient descent-driven all-in-one deep unfolding networks," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2025, pp. 7524–7533.

[55] X. Zhang, J. Ma, G. Wang, Q. Zhang, H. Zhang, and L. Zhang, "Perceive-IR: Learning to perceive degradation better for All-in-One image restoration," *IEEE Trans. Image Process.*, early access, May 9, 2025, doi: 10.1109/TIP.2025.3566300.

[56] C. Li, S. Anwar, and F. Porikli, "Underwater scene prior inspired deep underwater image and video enhancement," *Pattern Recognit.*, vol. 98, Feb. 2020, Art. no. 107038.

[57] Y. Wang, J. Guo, H. Gao, and H. Yue, "UIECˆ2-net: CNN-based underwater image enhancement using two color space," *Signal Process., Image Commun.*, vol. 96, Aug. 2021, Art. no. 116250.

[58] C. Li et al., "An underwater image enhancement benchmark dataset and beyond," *IEEE Trans. Image Process.*, vol. 29, pp. 4376–4389, 2020.

[59] H. Li, J. Li, and W. Wang, "A fusion adversarial underwater image enhancement network with a public test dataset," 2019, *arXiv:1906.06819*.

[60] Y.-T. Peng, K. Cao, and P. C. Cosman, "Generalization of the dark channel prior for single image restoration," *IEEE Trans. Image Process.*, vol. 27, no. 6, pp. 2856–2868, Jun. 2018.

[61] C. Fabbri, J. Islam, and J. Sattar, "Enhancing underwater imagery using generative adversarial networks," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, Jan. 2018, pp. 7159–7165.

[62] M. J. Islam, Y. Xia, and J. Sattar, "Fast underwater image enhancement for improved visual perception," *IEEE Robot. Autom. Lett.*, vol. 5, no. 2, pp. 3227–3234, Apr. 2020.

[63] C. Li, S. Anwar, J. Hou, R. Cong, C. Guo, and W. Ren, "Underwater image enhancement via medium transmission-guided multi-color space embedding," *IEEE Trans. Image Process.*, vol. 30, pp. 4985–5000, 2021.

[64] A. Hore and D. Ziou, "Image quality metrics: PSNR vs. SSIM," in *Proc. 20th Int. Conf. Pattern Recognit.*, Aug. 2010, pp. 2366–2369.

[65] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.

[66] X. Guo, Y. Li, and H. Ling, "LIME: Low-light image enhancement via illumination map estimation," *IEEE Trans. Image Process.*, vol. 26, no. 2, pp. 982–993, Feb. 2017.

[67] S. Wang, J. Zheng, H.-M. Hu, and B. Li, "Naturalness preserved enhancement algorithm for non-uniform illumination images," *IEEE Trans. Image Process.*, vol. 22, no. 9, pp. 3538–3548, Sep. 2013.

[68] R. Wang, Q. Zhang, C. Fu, X. Shen, W. Zheng, and J. Jia, "Underexposed photo enhancement using deep illumination estimation," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2019, pp. 6842–6850.

[69] W. Wang, W. Chen, W. Yang, and J. Liu, "GLADNet: Low-light enhancement network with global awareness," in *Proc. IEEE Int. Conf. Autom. Face Gesture Recognit.*, May 2018, pp. 751–755.

[70] Y. Jiang et al., "EnlightenGAN: Deep light enhancement without paired supervision," *IEEE Trans. Image Process.*, vol. 30, pp. 2340–2349, 2021.

[71] C. Guo et al., "Zero-reference deep curve estimation for low-light image enhancement," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2020, pp. 1780–1789.

[72] R. Liu, L. Ma, J. Zhang, X. Fan, and Z. Luo, "Retinex-inspired unrolling with cooperative prior architecture search for low-light image enhancement," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2021, pp. 10556–10565.

[73] C. Wei, W. Wang, W. Yang, and J. Liu, "Deep retinex decomposition for low-light enhancement," 2018, *arXiv:1808.04560*.

[74] Y. Zhang, J. Zhang, and X. Guo, "Kindling the darkness: A practical low-light image enhancer," in *Proc. ACM Int. Conf. Multimedia*, Jan. 2019, pp. 1632–1640.

[75] A. Ignatov, N. Kobyshev, R. Timofte, and K. Vanhoey, "DSLR-quality photos on mobile devices with deep convolutional networks," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 3277–3285.

[76] W. Yang, S. Wang, Y. Fang, Y. Wang, and J. Liu, "From fidelity to perceptual quality: A semi-supervised approach for low-light image enhancement," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2020, pp. 3063–3072.

[77] S. W. Zamir et al., "Learning enriched features for real image restoration and enhancement," in *Proc. Eur. Conf. Comput. Vis.*, vol. 12370, 2020, pp. 492–511.

[78] Y. Zhang, X. Guo, J. Ma, W. Liu, and J. Zhang, "Beyond brightening low-light images," *Int. J. Comput. Vis.*, vol. 129, no. 4, pp. 1013–1037, Apr. 2021.

[79] X. Lei, Z. Fei, W. Zhou, H. Zhou, and M. Fei, "Low-light image enhancement using the cell vibration model," *IEEE Trans. Multimedia*, vol. 25, pp. 4439–4454, 2022.

[80] L. Ma, T. Ma, R. Liu, X. Fan, and Z. Luo, "Toward fast, flexible, and robust low-light image enhancement," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2022, pp. 5637–5646.

[81] O. Özdenizci and R. Legenstein, "Restoring vision in adverse weather conditions with patch-based denoising diffusion models," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 45, no. 8, pp. 10346–10357, Aug. 2023.

[82] Y. Jin, W. Yan, W. Yang, and R. T. Tan, "Structure representation network and uncertainty feedback learning for dense non-uniform fog removal," in *Proc. Asian Conf. Comput. Vis.*, Jan. 2023, pp. 155–172.

[83] C. Guo, Q. Yan, S. Anwar, R. Cong, W. Ren, and C. Li, "Image dehazing transformer with transmission-aware 3D position embedding," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2022, pp. 5812–5820.

[84] Y. Song, Z. He, H. Qian, and X. Du, "Vision transformers for single image dehazing," *IEEE Trans. Image Process.*, vol. 32, pp. 1927–1941, 2023.

[85] Y. Yang, C. Wang, R. Liu, L. Zhang, X. Guo, and D. Tao, "Self-augmented unpaired image dehazing via density and depth decomposition," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2022, pp. 2037–2046.

[86] Z. Chen, Y. Wang, Y. Yang, and D. Liu, "PSD: Principled synthetic-to-real dehazing guided by physical priors," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2021, pp. 7180–7189.

[87] Z. Zheng et al., "Ultra-high-definition image dehazing via multi-guided bilateral learning," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2021, pp. 16180–16189.

[88] H. Dong et al., "Multi-scale boosted dehazing network with dense feature fusion," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jan. 2020, pp. 2157–2167.

[89] Y. Shao, L. Li, W. Ren, C. Gao, and N. Sang, "Domain adaptation for image dehazing," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2020, pp. 2808–2817.

[90] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014, *arXiv:1412.6980*.

[91] R. Zhang, P. Isola, A. A. Efros, E. Shechtman, and O. Wang, "The unreasonable effectiveness of deep features as a perceptual metric," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 586–595.

[92] A. Mittal, R. Soundararajan, and A. C. Bovik, "Making a 'completely blind' image quality analyzer," *IEEE Signal Process. Lett.*, vol. 20, no. 3, pp. 209–212, Mar. 2012.

[93] M. Yang and A. Sowmya, "An underwater color image quality evaluation metric," *IEEE Trans. Image Process.*, vol. 24, no. 12, pp. 6062–6071, Dec. 2015.

[94] K. He, J. Sun, and X. Tang, "Single image haze removal using dark channel prior," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 12, pp. 2341–2353, Dec. 2011.

[95] X. Liu, Y. Ma, Z. Shi, and J. Chen, "GridDehazeNet: Attention-based multi-scale network for image dehazing," in *Proc. IEEE/CVF Int. Conf. Comput. Vis.*, Oct. 2019, pp. 7313–7322.

[96] H. Li, R. Qiao, P. Yu, H. Li, and M. Tan, "CTHD-net: CNN-transformer hybrid dehazing network via residual global attention and gated boosting strategy," *J. Vis. Commun. Image Represent.*, vol. 99, Mar. 2024, Art. no. 104066.

[97] J. Zhang, Y. Cao, S. Fang, Y. Kang, and C. W. Chen, "Fast haze removal for nighttime image using maximum reflectance prior," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jul. 2017, pp. 7016–7024.

[98] C. Ancuti, C. O. Ancuti, C. De Vleeschouwer, and A. C. Bovik, "Nighttime dehazing by fusion," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Sep. 2016, pp. 2256–2260.

[99] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," in *Proc. IEEE Int. Conf. Comput. Vis.*, Oct. 2017, pp. 2223–2232.

[100] S. Rajagopalan and V. M. Patel, "AWRaCLe: all-weather image restoration using visual in-context learning," in *Proc. AAAI Conf. Artif. Intell.*, Aug. 2024, pp. 6675–6683.