

Degradation-Aware Dynamic Frequency Routing for Underwater Image Enhancement

Liheng Shi

石磊

Na Liu

liuna@dlpu.edu.cn

刘娜

Research Article

Keywords: underwater image enhancement, degradation awareness, frequency routing, spatial-frequency fusion, image restoration, lightweight network

Posted Date: October 14th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10486869/v1>

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

[Read Full License](#)

Additional Declarations: No competing interests reported.

Degradation-Aware Dynamic Frequency Routing for Underwater Image Enhancement

Liheng Shi¹ and Na Liu^{1*}

^{1*}School of Information Science and Engineering, Dalian Polytechnic University, Dalian, 116034, Liaoning, China.

*Corresponding author(s). E-mail(s): liuna@dlpu.edu.cn;

Abstract

Underwater images are jointly degraded by wavelength-selective absorption, scattering, and non-uniform illumination, and different scenes require different levels of color, structure, and texture restoration. Existing methods commonly rely on fixed enhancement paths or implicit feature modulation, limiting explicit coordination across frequency, spatial, and fusion domains. We therefore propose Degradation-Aware Dynamic Frequency Routing for Underwater Image Enhancement (DAFR-UIE). A Shared Degradation Embedding jointly conditions dynamic frequency routing, spatial modulation, and adaptive spatial-frequency fusion, enabling sample-dependent restoration of color, structure, and detail. On the Large Scale Underwater Image (LSUI) and Underwater Image Enhancement Benchmark (UIEB) datasets, DAFR-UIE achieves peak signal-to-noise ratio (PSNR)/structural similarity index measure (SSIM) values of 27.11 decibels (dB)/0.959 and 23.69 dB/0.897, respectively, with learned perceptual image patch similarity (LPIPS) values of 0.096 and 0.107 and only 2.02 million parameters. Ablations, severe-degradation analysis, and fixed-network interventions attribute the gains to the shared representation and input-dependent controls, while spectral and internal-representation analyses show that the control variables vary with degradation proxies. These results support input-conditioned resource allocation as a compact approach to underwater image enhancement.

Keywords: underwater image enhancement, degradation awareness, frequency routing, spatial-frequency fusion, image restoration, lightweight network

1 Introduction

Underwater image enhancement supports marine observation, underwater robotics, and ecological monitoring. Wavelength-selective absorption causes color casts, scattering reduces contrast and blurs edges, and non-uniform illumination produces spatially varying degradation[1–4]. Traditional methods use color correction, multi-scale fusion, or image-formation models[3, 5, 6], whereas deep methods employ convolutional

neural networks (CNNs), Transformers, diffusion models, and frequency-domain modeling[1, 4, 7–12]. Most enhancement policies nevertheless remain fixed or implicit in mixed features.

Different inputs require different frequency compensation: absorption and illumination mainly alter slowly varying color and brightness, whereas scattering and blur attenuate contours and fine structures. These effects are also spatially non-uniform because attenuation, turbidity, scene depth, and artificial lighting vary across scenes

and regions. A uniform policy may therefore over-enhance mild degradation or under-correct severe degradation. This motivates joint frequency allocation, spatial modulation, and feature fusion conditioned on the observed input.

We propose DAFR-UIE, in which a Shared Degradation Estimator provides a common state for a Dynamic Frequency Router, a Spatial Degradation Map, and an Adaptive Fusion Gate. This within-block design couples image-level allocation with local residual modulation rather than estimating each control independently. We evaluate restoration quality, cross-dataset transfer, component contributions, difficult cases, and learned-control behavior under paired and fixed-network protocols.

The main contributions are as follows:

- (1) We introduce a within-block shared degradation representation that provides a common conditioning state for frequency, spatial, and fusion control.
- (2) We develop a competitive frequency-routing and spatial-frequency fusion mechanism for sample-dependent restoration.
- (3) We validate the proposed controls through paired and cross-dataset evaluation, retrained ablations, severe-degradation analysis, and fixed-network interventions.

2 Related Work

2.1 Traditional Underwater Image Enhancement

Traditional methods mainly estimate background light and transmission through white balancing, multi-scale fusion, or physical models[3, 5, 6]. Physical models associate color attenuation and scattering with interpretable parameters, whereas multi-scale fusion combines color, contrast, and saliency cues without requiring explicit imaging parameters. In real water bodies, however, spectral attenuation, suspended particles, and artificial illumination often vary jointly. Fixed priors may therefore leave residual color casts under model mismatch, while uniform fusion rules may over-enhance local regions. These methods offer clear modeling foundations but limited sample-level adaptability to complex degradation combinations.

2.2 Deep Learning-Based Underwater Image Enhancement

End-to-end methods have evolved from generative learning[2, 13] to color-space modeling[4], Transformers[7, 11, 12], contrastive learning[14], uncertainty modeling[15], diffusion models[16], and spatial-frequency collaboration[8–10]. These models improve nonlinear restoration capacity, but degradation adaptation is usually implicit in their features or branches. Variations in underwater imaging conditions and data distributions also widen the gap between training and test domains, and domain-adaptation studies have shown that a fixed mapping alone cannot cover all real scenes[17]. We therefore investigate whether an observable shared variable can jointly regulate frequency resources, spatial strength, and fusion ratios rather than compressing all degradation differences into a static enhancement rule.

2.3 Frequency-Domain Enhancement and Dynamic Routing

Frequency-domain methods separate global color from local structure[9], but fixed band weights can over-enhance details or under-correct color. Dynamic processing instead allocates computation according to the input, although its routing decisions must correspond to meaningful image components and spatially varying degradation. These limitations motivate degradation-conditioned frequency allocation coordinated with spatially varying restoration.

3 Method

3.1 Overall Framework

The overall architecture of DAFR-UIE follows an encoder–memory–decoder design, as illustrated in Fig. 1. The backbone extracts shallow and bottleneck features from $I \in \mathbb{R}^{3 \times H \times W}$:

$$F_1 = \phi_1(I), \quad F_2 = \mathcal{P}(\phi_2(F_1)), \quad F_b = \mathcal{P}(\phi_3(F_2)). \quad (1)$$

where ϕ_i is a convolutional layer and $\mathcal{P}$ is max pooling. Memory blocks composed of DAFR-Blocks aggregate long- and short-term features

using gated 1×1 convolutions. Their weighted outputs enter a bilinear-upsampling decoder, which reconstructs the image in residual form:

$$\hat{I} = \psi(F_{\text{dec}}) + \lambda_r I, \quad (2)$$

where ψ is the reconstruction convolution and the implemented residual scale is $\lambda_r = 0.1$. Auxiliary decoder outputs use the same form for multi-scale supervision. The default model uses 16 base channels, four memory blocks, and four DAFR units per block. This compact configuration facilitates evaluation of the conditional-control mechanism without relying on excessive model capacity.

3.2 Degradation-Aware Estimator

The Shared Degradation Estimator applies two convolution-group-normalization (GroupNorm)-sigmoid-linear-unit (SiLU) layers to X (Fig. 2):

$$D = E_\theta(X). \quad (3)$$

Three lightweight heads then produce the Spatial Degradation Map, routing logits, and Adaptive Fusion Gate; the routing head uses global average pooling (GAP):

$$M_s = \sigma(h_s(D)), \mathbf{a} = h_r(\text{GAP}(D)), G_f = \sigma(h_f(D)). \quad (4)$$

Here, $M_s \in [0, 1]^{1 \times H \times W}$ is a spatial degradation map, $\mathbf{a} \in \mathbb{R}^3$ contains low/mid/high-frequency logits, and $G_f \in [0, 1]^{C \times H \times W}$ controls residual injection. The spatial and fusion heads use 1×1 convolution and sigmoid; the routing head uses global average pooling and two 1×1 convolutions. All are readouts of the same embedding.

3.3 Dynamic Frequency Routing Module

The Dynamic Frequency Router allocates enhancement among frequency subspaces. It applies a two-dimensional (2D) fast Fourier transform (FFT), specifically a real FFT (RFFT), to X_f :

$$\mathcal{F}(X_f) = \text{RFFT2}(X_f). \quad (5)$$

Fixed radial masks define low ($r \leq 0.16$), mid ($0.16 < r \leq 0.34$), and high ($r > 0.34$) bands,

and the inverse real FFT (IRFFT) returns them to the spatial domain:

$$B_k = \text{IRFFT2}(\mathcal{F}(X_f) \odot \Omega_k), \quad k \in \{l, m, h\}. \quad (6)$$

where Ω_k is the mask for band k . Low frequencies capture color and illumination, mid frequencies emphasize contours and contrast, and high frequencies represent fine details and noise. A temperature-scaled softmax converts logits to weights:

$$\mathbf{w} = \text{Softmax}(\mathbf{a}/\tau) = [w_l, w_m, w_h], \quad \sum_k w_k = 1. \quad (7)$$

In the full model, mid and high frequencies are combined, giving

$$\tilde{w}_l = w_l, \quad \tilde{w}_h = w_m + w_h. \quad (8)$$

Disabled branches receive zero weight and the remainder is renormalized. The enhanced feature is

$$F_{\text{freq}} = \tilde{w}_l R_l(B_l) + \tilde{w}_h \cdot \frac{R_m(B_m) + R_h(B_h)}{2}, \quad (9)$$

where R_k is a depthwise-separable refinement branch. Decomposition is performed on learned features rather than red-green-blue (RGB) pixels. The input-dependent routing weights suppress unnecessary structural amplification in mildly degraded or noisy regions.

3.4 Spatial-Frequency Collaborative Fusion

Adaptive Spatial-Frequency Enhancement determines where and how strongly routed frequency information is applied. A 1×1 convolution splits the input into X_s and X_f . A visual state-space MambaLayer with 2D selective scan, derived from Mamba and VMamba[18, 19], models spatial context, and M_s modulates its residual:

$$F_{\text{sp}} = X_s + M_s \odot \Delta_s(X_s). \quad (10)$$

where Δ_s is the learned enhancement residual. The frequency branch computes

$$\Delta_f = F_{\text{freq}} - X_f. \quad (11)$$

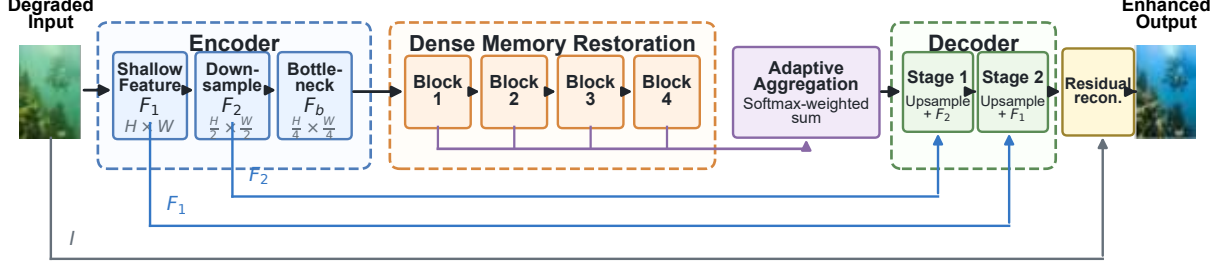


Fig. 1 Macro architecture of DAFR-UIE. The encoder extracts multiscale features F_1 , F_2 , and F_b ; four MemoryBlockDAFR stages perform dense restoration; adaptive memory aggregation combines their outputs; and the decoder reconstructs the enhanced image through feature and image residual connections

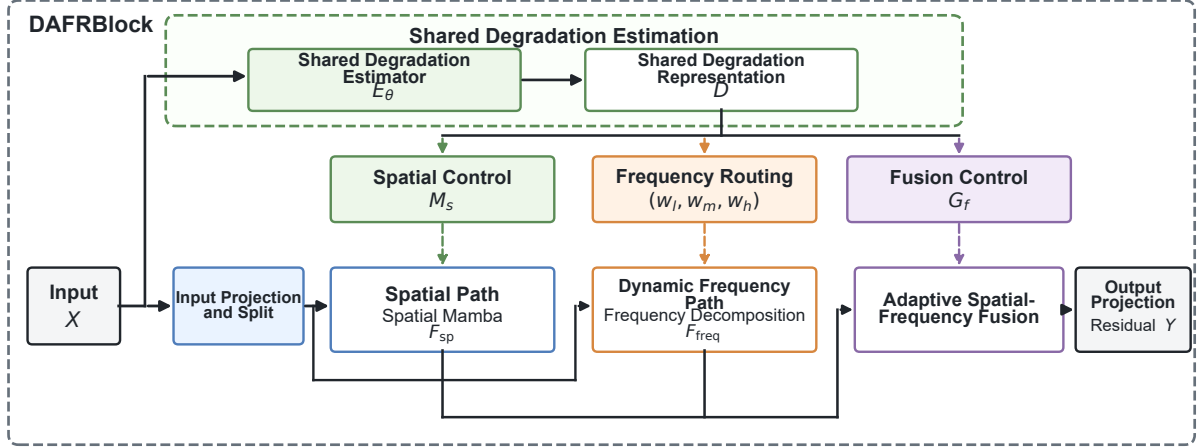


Fig. 2 Simplified architecture of the DAFRBlock. A within-block Shared Degradation Representation generates the spatial control M_s , frequency-routing weights (w_l, w_m, w_h) , and fusion control G_f , which regulate spatial enhancement, frequency allocation, and adaptive spatial-frequency fusion, respectively. Detailed operations are given in Eqs. (3)–(13)

The same map aligns frequency-residual strength with local degradation:

$$R_f = \beta(0.5 + M_s) \odot \Delta_f. \quad (12)$$

The Adaptive Fusion Gate is biased by the combined mid/high routing weight and then controls the final injection:

$$\begin{aligned} \tilde{G}_f &= \text{clip}_{[0,1]}(G_f(0.5 + \tilde{w}_h)), \\ F_{\text{out}} &= F_{\text{sp}} + \alpha \tilde{G}_f \odot R_f. \end{aligned} \quad (13)$$

where the implementation bounds the learnable scales as $\alpha = 0.5 + 0.5\sigma(a_\alpha)$ and $\beta = 0.35 + 0.65\sigma(a_\beta)$.

3.5 Design Rationale and Control Mechanism

The softmax normalization in Eq. (7) imposes competition among the frequency bands under a fixed allocation budget, while the temperature τ controls the smoothness of the routing distribution. All three controls are derived from the shared representation D but operate at different scopes. Global average pooling produces image-level frequency allocation, the full-resolution spatial map modulates locally varying degradation, and the channel-spatial gate regulates residual injection. This global-allocation–local-modulation–fusion chain provides coordinated control within each DAFRBlock while keeping the conditioning structure compact.

3.6 Loss Function

The objective combines pixel, structural, frequency, chromatic, detail, auxiliary, and control terms:

$$\begin{aligned}\mathcal{L} = & \lambda_1\mathcal{L}_1 + \lambda_2\mathcal{L}_{mse} + \lambda_3\mathcal{L}_{ssim} \\ & + \lambda_4\mathcal{L}_{fdl} + \lambda_5\mathcal{L}_{chroma} + \lambda_6\mathcal{L}_{detail} \\ & + \lambda_7\mathcal{L}_{aux} + \lambda_8\mathcal{L}_{ctrl}.\end{aligned}\quad (14)$$

$\mathcal{L}_1$ and the mean-squared-error (MSE) term $\mathcal{L}_{mse}$ supervise pixels; $\mathcal{L}_{ssim}$ follows SSIM[20], and $\mathcal{L}_{fdl}$ adapts focal frequency loss[21]. The custom $\mathcal{L}_{chroma}$ and $\mathcal{L}_{detail}$ terms constrain Commission Internationale de l’Éclairage (CIE) L*a*b* chromaticity and gradients/local contrast, respectively, while $\mathcal{L}_{aux}$ supervises multi-scale outputs. $\mathcal{L}_{ctrl}$ links the three controls to input-reference residual proxies using band-wise Kullback–Leibler (KL) divergence, Smooth L1, high-frequency degradation maps, and minimum-variance regularization.

The fixed progressive schedule, shared by all comparisons, first stabilizes coarse reconstruction and subsequently emphasizes structural, chromatic, and pixel accuracy. The control loss uses soft residual proxies because underwater degradations are continuous and overlapping rather than mutually exclusive categories.

4 Experiments

4.1 Datasets and Experimental Settings

Paired experiments use the LSUI dataset[7] (3,879 training pairs and 400 test pairs) and UIEB dataset[1] (800 training pairs and 90 test pairs). LSUI is the principal benchmark for quantitative and mechanism analysis, while UIEB tests paired reconstruction under different image statistics. The training schedule and model-selection rule are fixed in advance; all main results use the exponential-moving-average (EMA) checkpoint at epoch 250, and test sets are used only for final evaluation.

Cross-dataset experiments use the Underwater Image Quality Set (UIQS) of the Real-world Underwater Image Enhancement (RUIE) benchmark[22], denoted RUIE-UIQS, and the U45

test set[23]. Both contain unpaired real scenes with varied color casts, contrast, and acquisition conditions. Trained weights are applied directly without target-domain retraining or parameter selection. Paired evaluation measures reference fidelity, whereas the underwater image quality measure (UIQM) and underwater color image quality evaluation (UCIQE) assess transfer when no unique target is available.

Training uses 256×256 inputs, random cropping, horizontal flipping, seed 2026, and 250 epochs. The Adam optimizer[24] uses $\beta_1 = 0.5$, with learning-rate warmup followed by cosine annealing; EMA model weights[25] are used for evaluation. Speed results compare ablations under identical hardware and implementation conditions.

4.2 Evaluation Metrics

We use PSNR and SSIM for reference-based fidelity[20, 26], FID and LPIPS for distributional and perceptual quality[27, 28], and UIQM and UCIQE for no-reference underwater image assessment[29, 30]. Parameter count and latency characterize model efficiency. All metrics except FID are computed per image and then averaged. Retrained ablations evaluate component contributions, whereas inference-time interventions examine learned controls with fixed network weights.

4.3 Comparison on the Paired LSUI and UIEB Datasets

Table 1 compares public metrics and parameter counts on LSUI and UIEB.

On LSUI, DAFR-UIE achieves the best reported SSIM and LPIPS in Table 1 using only 2.02 million parameters. Although SS-UIE obtains a higher PSNR, DAFR-UIE provides a stronger structure-perception-complexity balance.

On UIEB, DAFR-UIE maintains the lowest reported LPIPS and competitive PSNR/SSIM, indicating consistent perceptual restoration across datasets.

4.4 No-Reference Generalization on RUIE-UIQS and U45

Under the direct cross-dataset protocol described in Sect. 4.1, Table 2 reports UIQM and UCIQE on RUIE-UIQS and U45.

Table 1 Quantitative comparison on the LSUI and UIEB test sets; published comparison values are from the original sources [7–12, 14, 16]

(a) LSUI						(b) UIEB					
Method	FID↓	LPIPS↓	PSNR↑	SSIM↑	Params/M ↓	Method	FID↓	LPIPS↓	PSNR↑	SSIM↑	Params/M ↓
DAFR-UIE	26.33	0.096	27.11	0.959	2.02	DAFR-UIE	28.25	0.107	23.69	0.897	2.02
SS-UIE [8]	25.98	0.106	28.87	0.952	4.25	SS-UIE [8]	27.31	0.117	24.19	0.897	4.25
UIEFormer [11]	–	–	–	–	1.70	UIEFormer [11]	–	–	24.78	0.856	1.70
U-TransCNN [12]	–	–	–	–	15.38	U-TransCNN [12]	–	–	24.47	0.876	15.38
WF-Diff [9]	26.75	0.110	27.26	0.944	43.72	WF-Diff [9]	27.85	0.125	23.86	0.873	43.72
HCLR-Net [10]	–	–	–	–	4.65	HCLR-Net [10]	–	–	22.59	0.898	4.65
CCL-Net [14]	–	–	–	–	0.55	CCL-Net [14]	–	–	20.18	0.866	0.55
DM-Water [16]	27.91	0.114	27.65	0.887	10.13	DM-Water [16]	31.07	0.144	21.88	0.819	10.13
U-Shape [7]	28.56	0.103	24.16	0.932	30.13	U-Shape [7]	46.11	0.226	21.25	0.845	30.13

Note: “–” denotes that no reliable public result was found. Every comparison method is linked to its original paper by the citation following the method name; metrics and parameter counts follow the cited papers or their public implementations. Bold indicates the best reported value shown for each dataset.

Table 2 No-reference generalization results on RUIE-UIQS and U45; published comparison values are from the original sources [4, 6–9, 16, 22, 31–34]

Method	RUIE UIQM↑	RUIE UCIQE↑	U45 UIQM↑	U45 UCIQE↑
DAFR-UIE	3.240	0.449	3.552	0.636
SS-UIE [8]	–	–	3.246	0.703
WF-Diff [9]	–	–	3.181	0.619
DM-Water [16]	–	–	3.086	0.634
U-Shape [7]	–	–	3.151	0.592
MLLE [6]	–	–	2.683	0.456
U-Color [4]	–	–	3.104	0.586
Input [22]	0.337	0.290	–	–
DPATN [31]	3.754	0.480	–	–
UHP [32]	3.109	0.503	–	–
HLPcb [33]	2.968	0.477	–	–
CAPcb [34]	3.288	0.471	–	–

Note: “–” denotes an unreported result. RUIE-UIQS classical-baseline values and the color-balanced variants denoted by “cb” follow the RUIE benchmark protocol[22]; other values are from the cited papers or public tables. Bold indicates the best value for each metric.

DAFR-UIE substantially improves both no-reference metrics over the raw RUIE-UIQS inputs and achieves the highest UIQM on U45. Its UCIQE is competitive but lower than SS-UIE, indicating a favorable overall quality balance rather than optimization toward a single index.

4.5 Ablation Study

Table 3 ablates frequency branches, the Shared Degradation Embedding, Frequency Decomposition, the Adaptive Fusion Gate, and all degradation-aware controls under one protocol. Differences use Full as reference. In the w/o Frequency Decomposition setting, the frequency branch retains a comparable spatial refinement path but removes explicit FFT-based band separation.

Removing the degradation embedding or all degradation-aware controls causes the largest performance degradation, confirming that the gains arise primarily from input conditioning rather than branch capacity alone. Frequency-branch ablations produce the next-largest drops, whereas frequency decomposition and fusion gating provide smaller but consistent structural improvements.

4.6 Severe-Degradation Subset Analysis

To evaluate difficult cases, we test the main ablation settings on the severe-degradation subset of LSUI (Table 4).

The degradation embedding becomes more important on the severe subset, where its removal causes a larger PSNR reduction than on the full test set. The increased sensitivity to the low-frequency path and explicit decomposition is consistent with stronger color and illumination distortion in difficult scenes.

4.7 Mechanism Intervention and Interpretability

We complement retrained ablations by fixing or shuffling the three controls in the trained full model (Table 5).

Uniform routing reduces PSNR by 2.342 dB, showing that competitive band allocation matters, while shuffled routing tests whether valid weights remain matched to the corresponding input. Shuffling the spatial map causes the largest reduction (10.124 dB) and lowers Spatial Spearman to 0.041, associating local modulation with the current

Table 3 Main ablation results for DAFR-UIE
(a) Restoration quality

Setting	PSNR (dB)	SSIM	Δ PSNR	Δ SSIM
	$\uparrow$	$\uparrow$	(dB)	
Full	27.11	0.959	0.000	0.000
w/o Frequency Branch	26.57	0.948	-0.540	-0.011
w/o Low-frequency Branch	26.72	0.947	-0.390	-0.012
w/o High-frequency Branch	26.62	0.946	-0.496	-0.013
w/o Degradation Embedding	26.06	0.946	-1.050	-0.013
w/o Frequency Decomposition	26.96	0.950	-0.154	-0.009
w/o Fusion Gate	27.07	0.951	-0.041	-0.009
w/o Degradation-aware Control	26.00	0.947	-1.111	-0.012

Note: LSUI test set at 256×256 input resolution. Δ PSNR and Δ SSIM are differences from Full. Latency was measured on one NVIDIA RTX 4090.

Table 4 Ablation results on the severe-degradation subset of LSUI ($N = 133$)

Setting	PSNR (dB)	Δ PSNR	SSIM	Δ SSIM
	$\uparrow$	(dB)	$\uparrow$	
Full	26.33	0.000	0.950	0.000
w/o Frequency Branch	25.78	-0.55	0.941	-0.009
w/o Low-frequency Branch	25.82	-0.51	0.940	-0.010
w/o High-frequency Branch	26.10	-0.23	0.940	-0.010
w/o Degradation Embedding	24.92	-1.41	0.936	-0.014
w/o Frequency Decomposition	25.87	-0.46	0.942	-0.008
w/o Fusion Gate	26.01	-0.32	0.941	-0.009
w/o Degradation-aware Control	25.54	-0.79	0.941	-0.009

Note: Δ PSNR and Δ SSIM are differences from Full.

Table 5 Mechanism-intervention results for DAFR-UIE

Intervention	Restoration effect		Mechanism diagnostics		
	PSNR $\uparrow$	SSIM $\uparrow$	Routing	Spatial	Gate
	(Δ PSNR)	(Δ SSIM)	entropy	Spearman	mean
<i>none</i>	27.11 (0.000)	0.959 (0.000)	1.075	0.472	0.645
Uniform route	24.77 (-2.342)	0.951 (-0.008)	1.099	0.467	0.597
Fixed spatial	19.41 (-7.699)	0.908 (-0.051)	1.079	0.000	0.640
Fixed fusion	26.46 (-0.657)	0.958 (-0.001)	1.075	0.485	0.630
Shuffled route	26.24 (-0.870)	0.957 (-0.002)	1.075	0.464	0.644
Shuffled spatial	16.99 (-10.124)	0.838 (-0.121)	1.071	0.041	0.648
Shuffled fusion	26.92 (-0.191)	0.958 (-0.001)	1.076	0.483	0.643

Note: $N = 400$ LSUI test images per intervention. In the PSNR and SSIM columns, parenthesized second-line values are paired changes relative to *none*. Spatial Spearman is the rank correlation between the spatial map and the input-reference degradation proxy.

image rather than generic attention. Fusion-gate interventions have smaller effects (at most 0.657 dB), consistent with downstream residual adjustment.

4.8 Visual Comparison

Figure 3 shows mild, moderate, and severe LSUI cases. DAFR-UIE corrects residual casts in mild degradation and raises brightness, contrast, and texture visibility in harder cases. Errors remain concentrated around boundaries, highlights, and

(b) Model efficiency

Setting	Params (M)	Latency (ms)
	$\downarrow$	$\downarrow$
Full	2.02	43.41
w/o Frequency Branch	1.67	33.01
w/o Low-frequency Branch	2.02	51.84
w/o High-frequency Branch	2.02	48.14
w/o Degradation Embedding	1.52	43.80
w/o Frequency Decomposition	1.79	43.19
w/o Fusion Gate	1.99	50.35
w/o Degradation-aware Control	1.52	41.82

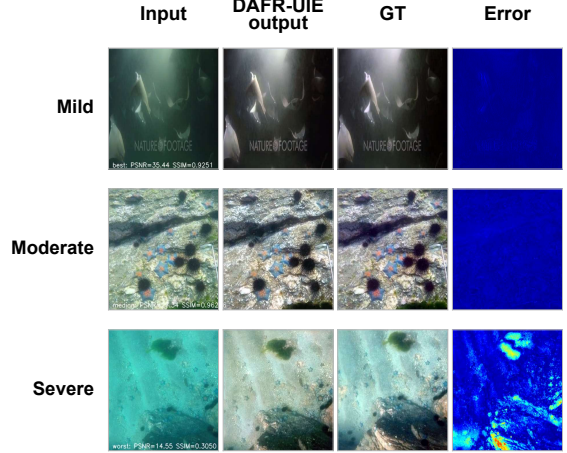


Fig. 3 Representative LSUI restoration results. Rows show mild, moderate, and severe cases; columns show the input, DAFR-UIE output, ground truth, and absolute error map

fine textures, and increase with degradation severity, consistent with the subset analysis.

4.9 Frequency Energy and Dynamic-Routing Interpretability

For 400 LSUI pairs, residual power is measured in low ($r \leq 0.16$), mid ($0.16 < r \leq 0.34$), and high ($r > 0.34$) bands and stratified by Input PSNR. Low-frequency energy dominates and varies with degradation (Fig. 4a).

After radial-frequency-squared prewhitening, mid/high components still differ across degradation levels (Kruskal-Wallis $p = 1.29 \times 10^{-5}$, 1.68×10^{-4} ; Fig. 4b). The variation therefore exceeds the natural low-frequency concentration of image energy and motivates an input-dependent rather than constant branch mixture.

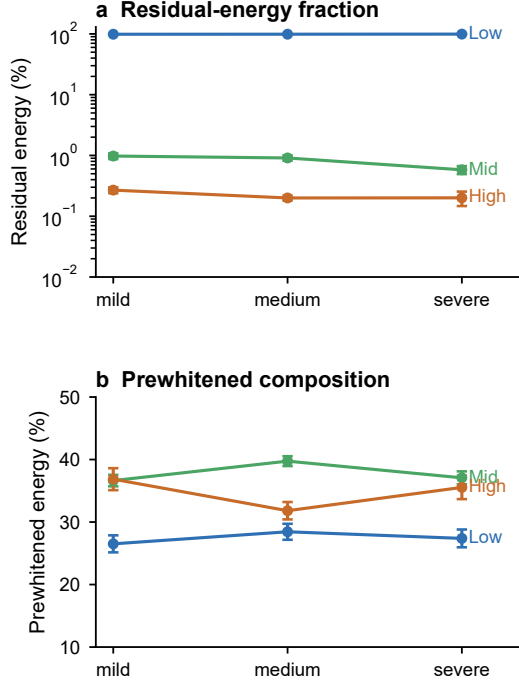


Fig. 4 LSUI residual-frequency composition: a, total fractions; b, prewhitened fractions. Error bars denote 95% CIs; $n = 133$ or 134 per group

Mean low/mid/high weights are 0.239/0.370/0.391 on LSUI and 0.505/0.271/0.224 on UIEB (Fig. 5). UIEB favors broad color/illumination correction, whereas LSUI allocates more weight to mid/high restoration. Within-dataset dispersion confirms that coefficients are sample dependent rather than dataset constants.

A principal component analysis (PCA) shows that the first two principal components (PCs) of the embedding explain 75.1% and 6.8% of variance, with centers shifting continuously along the first principal component (PC1) across input-PSNR strata (Fig. 6). This continuous structure matches gradual, mixed underwater degradation and supports continuous conditioning rather than discrete classes.

The spatial map responds strongly at edges, dark regions, highlights, and complex textures but weakly in smooth regions (Fig. 7), concentrating restoration where visibility or structure is locally compromised.

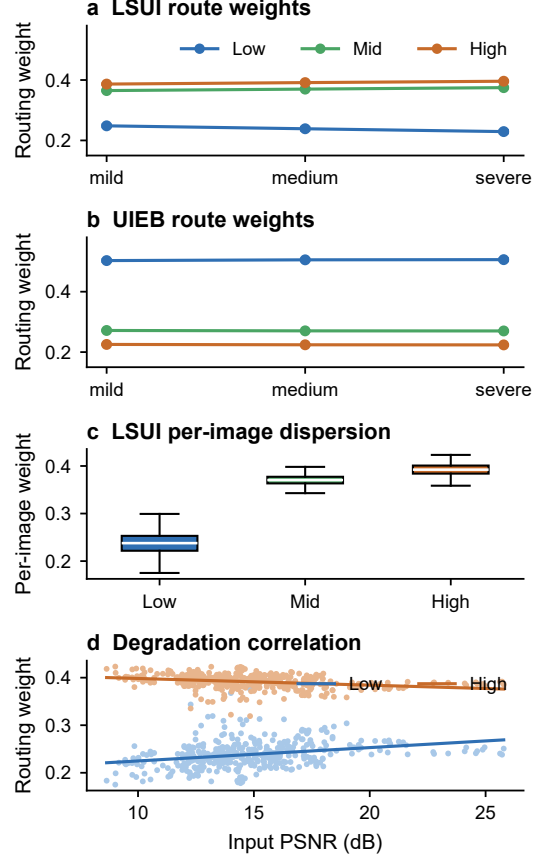


Fig. 5 Dynamic routing statistics: a–b, stratum means; c, per-image distributions; d, routing versus input PSNR. Error bars denote 95% CIs

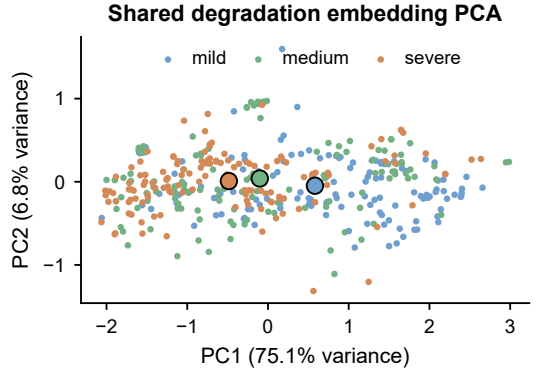


Fig. 6 PCA of the Shared Degradation Embedding

Representative coefficients depart from uniform allocation and agree with the aggregate routing statistics (Fig. 8).

These analyses jointly support degradation-dependent global routing and local modulation.

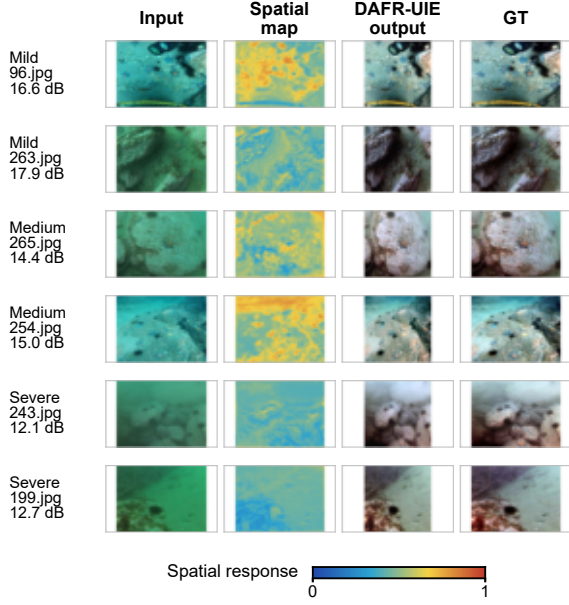


Fig. 7 Spatial Degradation Map

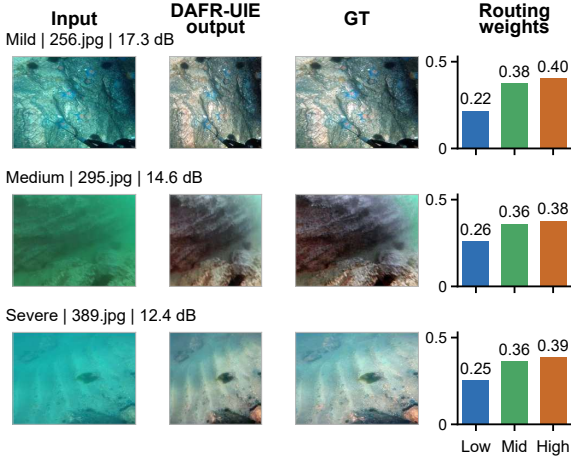


Fig. 8 Representative dynamic routing weights

5 Discussion

The results support viewing underwater enhancement as an input-conditioned allocation problem rather than a fixed restoration mapping. The Shared Degradation Embedding provides a common degradation state, while frequency routing, spatial modulation, and fusion gating operate at complementary global, local, and downstream levels. Their unequal intervention effects suggest that spatial localization is the dominant control, frequency routing determines spectral emphasis, and fusion gating mainly refines the resulting residual.

Cross-dataset results indicate transfer without fine-tuning, but the current model has several limitations. Fixed radial boundaries may not be optimal across cameras, water types, artificial lighting, or motion blur, and frame-wise routing may introduce temporal fluctuations in videos. Future work will investigate learnable ordered partitions, physics-guided conditioning, temporal routing, subjective evaluation, and downstream detection and robotic perception tasks.

6 Conclusion

DAFR-UIE introduces a shared degradation-conditioned framework for dynamic frequency allocation and spatial-frequency restoration. With 2.02 million parameters, it achieves competitive paired and cross-dataset performance. Ablations, difficult-case tests, interventions, and internal analyses support its compact, sample-dependent design within the present scope of static images and fixed radial frequency partitions.

Statements and Declarations

Funding. This work was supported by the Liaoning Provincial Department of Education General Project under Grant JYTMS20230427.

Competing Interests. The authors declare no competing interests.

Author Contributions. L.S. contributed to conceptualization, methodology, software, experiments, data analysis, visualization, and manuscript preparation. N.L. contributed to supervision, validation, and manuscript review. All authors read and approved the final manuscript.

Data Availability. The LSUI, UIEB, RUIE-UIQS, and U45 datasets used in this study are publicly available from their original repositories, as linked in Sect. 4.1. No new dataset was created.

Code Availability. The implementation code and trained weights are available from the corresponding author upon reasonable request.

Ethics Approval. Not applicable.

Consent to Participate. Not applicable.

Consent for Publication. Not applicable.

References

- [1] C. Li, C. Guo, W. Ren, R. Cong, J. Hou, S. Kwong, D. Tao. An Underwater Image Enhancement Benchmark Dataset and Beyond. *IEEE Trans. Image Process.*, vol. 29, pp. 4376–4389, 2020. <https://doi.org/10.1109/TIP.2019.2955241>
- [2] J. Li, K. A. Skinner, R. M. Eustice, M. Johnson-Roberson. WaterGAN: Unsupervised Generative Network to Enable Real-time Color Correction of Monocular Underwater Images. *IEEE Robot. Autom. Lett.*, vol. 3, no. 1, pp. 387–394, 2018. <https://doi.org/10.1109/LRA.2017.2730363>
- [3] C. Li, S. Anwar, F. Porikli. Underwater Scene Prior Inspired Deep Underwater Image and Video Enhancement. *Pattern Recognit.*, vol. 98, art. 107038, 2020. <https://doi.org/10.1016/j.patcog.2019.107038>
- [4] C. Li, S. Anwar, J. Hou, R. Cong, C. Guo, W. Ren. Underwater Image Enhancement via Medium Transmission-Guided Multi-Color Space Embedding. *IEEE Trans. Image Process.*, vol. 30, pp. 4985–5000, 2021. <https://doi.org/10.1109/TIP.2021.3076367>
- [5] C. Ancuti, C. O. Ancuti, T. Haber, P. Bekaert. Enhancing Underwater Images and Videos by Fusion. In: *Proc. IEEE CVPR*, pp. 81–88, 2012. <https://doi.org/10.1109/CVPR.2012.6247661>
- [6] W. Zhang, P. Zhuang, H.-H. Sun, G. Li, S. Kwong, C. Li. Underwater Image Enhancement via Minimal Color Loss and Locally Adaptive Contrast Enhancement. *IEEE Trans. Image Process.*, vol. 31, pp. 3997–4010, 2022. <https://doi.org/10.1109/TIP.2022.3177129>
- [7] L. Peng, C. Zhu, L. Bian. U-Shape Transformer for Underwater Image Enhancement. *IEEE Trans. Image Process.*, vol. 32, pp. 3066–3079, 2023. <https://doi.org/10.1109/TIP.2023.3276332>
- [8] L. Peng, L. Bian. Adaptive Dual-domain Learning for Underwater Image Enhancement. *Proc. AAAI Conf. Artif. Intell.*, vol. 39, no. 6, pp. 6461–6469, 2025. <https://doi.org/10.1609/aaai.v39i6.32692>
- [9] C. Zhao, W. Cai, C. Dong, C. Hu. Wavelet-based Fourier Information Interaction with Frequency Diffusion Adjustment for Underwater Image Restoration. In: *Proc. IEEE/CVF CVPR*, pp. 8281–8291, 2024. <https://doi.org/10.1109/CVPR52733.2024.00791>
- [10] J. Zhou, J. Sun, C. Li, Q. Jiang, M. Zhou, K.-M. Lam, W. Zhang, X. Fu. HCLR-Net: Hybrid Contrastive Learning Regularization with Locally Randomized Perturbation for Underwater Image Enhancement. *Int. J. Comput. Vis.*, vol. 132, no. 10, pp. 4132–4156, 2024. <https://doi.org/10.1007/s11263-024-01987-y>
- [11] J. Qu, X. Cao, S. Jiang, J. You, Z. Yu. UIFormer: Lightweight Vision Transformer for Underwater Image Enhancement. *IEEE J. Oceanic Eng.*, vol. 50, no. 2, pp. 851–865, 2025. <https://doi.org/10.1109/JOE.2024.3519681>
- [12] H. Yao, R. Guo, Z. Zhao, Y. Zang, X. Zhao, T. Lei, H. Wang. U-TransCNN: A U-shape Transformer-CNN Fusion Model for Underwater Image Enhancement. *Displays*, vol. 88, art. 103047, 2025. <https://doi.org/10.1016/j.displa.2025.103047>
- [13] C. Fabbri, M. J. Islam, J. Sattar. Enhancing Underwater Imagery Using Generative Adversarial Networks. In: *Proc. IEEE ICRA*, pp. 7159–7165, 2018. <https://doi.org/10.1109/ICRA.2018.8460552>
- [14] Y. Liu, Q. Jiang, X. Wang, T. Luo, J. Zhou. Underwater Image Enhancement With Cascaded Contrastive Learning. *IEEE Trans. Multimedia*, vol. 27, pp. 1512–1525, 2025. <https://doi.org/10.1109/TMM.2024.3521739>
- [15] Z. Fu, W. Wang, Y. Huang, X. Ding, K.-K. Ma. Uncertainty Inspired Underwater Image Enhancement. In: *Computer Vision—ECCV 2022, LNCS vol. 13678*, pp. 465–482, Springer, 2022. https://doi.org/10.1007/978-3-031-19797-0_27
- [16] Y. Tang, H. Kawasaki, T. Iwaguchi. Underwater Image Enhancement by Transformer-based Diffusion Model with Non-uniform Sampling for Skip Strategy. In: *Proc. ACM Multimedia*, pp. 5419–5427, 2023. <https://doi.org/10.1145/3581783.3612378>
- [17] Z. Wang, L. Shen, M. Xu, M. Yu, K. Wang, Y. Lin. Domain Adaptation for Underwater Image Enhancement. *IEEE Trans. Image Process.*, vol. 32, pp. 1442–1457, 2023. <https://doi.org/10.1109/TIP.2023.3244647>
- [18] A. Gu, T. Dao. Mamba: Linear-Time Sequence Modeling with Selective State Spaces. *arXiv:2312.00752*, 2023. <https://doi.org/10.48550/arXiv.2312.00752>
- [19] Y. Liu, Y. Tian, Y. Zhao, H. Yu, L. Xie, Y. Wang, Q. Ye, J. Jiao, Y. Liu. VMamba: Visual State Space Model. *Adv. Neural Inf. Process. Syst.*, vol. 37, 2024. <https://doi.org/10.52202/079017-3273>
- [20] Z. Wang, A. C. Bovik, H. R. Sheikh, E. P. Simoncelli. Image Quality Assessment: From Error Visibility to Structural Similarity. *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, 2004. <https://doi.org/10.1109/TIP.2003.819861>
- [21] L. Jiang, B. Dai, W. Wu, C. C. Loy. Focal Frequency Loss for Image Reconstruction and Synthesis. In: *Proc. IEEE/CVF ICCV*, pp. 13919–13929, 2021. <https://doi.org/10.1109/ICCV48922.2021.01366>
- [22] R. Liu, X. Fan, M. Zhu, M. Hou, Z. Luo. Real-World Underwater Enhancement: Challenges, Benchmarks, and Solutions Under Natural Light. *IEEE Trans. Circuits Syst. Video Technol.*, vol. 30, no. 12, pp. 4861–4875, 2020. <https://doi.org/10.1109/TCSVT.2019.2963772>
- [23] H. Li, J. Li, W. Wang. A Fusion Adversarial Underwater Image Enhancement Network with a Public Test Dataset. *arXiv:1906.06819*, 2019. <https://doi.org/10.48550/arXiv.1906.06819>
- [24] D. P. Kingma, J. Ba. Adam: A Method for Stochastic Optimization. In: *Proc. International Conference on Learning Representations*, 2015. <https://doi.org/10.48550/arXiv.1412.6980>
- [25] A. Tarvainen, H. Valpola. Mean Teachers Are Better Role Models: Weight-Averaged Consistency Targets Improve Semi-Supervised Deep Learning Results. *Adv. Neural Inf. Process. Syst.*, vol. 30, 2017. <https://doi.org/10.48550/arXiv.1703.01780>
- [26] A. Hore, D. Ziou. Image Quality Metrics: PSNR vs. SSIM. In: *Proc. IEEE International Conference on Pattern Recognition*, pp. 2366–2369, 2010. <https://doi.org/10.1109/ICPR.2010.579>
- [27] M. Heusel, H. Ramsauer, T. Unterthiner, B. Nessler, S. Hochreiter. GANs Trained by a Two Time-Scale Update Rule Converge to a Local Nash Equilibrium. *Adv. Neural Inf. Process. Syst.*, vol. 30, 2017. <https://doi.org/10.48550/arXiv.1706.08500>
- [28] R. Zhang, P. Isola, A. A. Efros, E. Shechtman, O. Wang. The Unreasonable Effectiveness of Deep Features as a Perceptual Metric. In: *Proc. IEEE CVPR*, pp. 586–595, 2018. <https://doi.org/10.1109/CVPR.2018.00068>
- [29] K. Panetta, C. Gao, S. Agaian. Human-Visual-System-Inspired Underwater Image Quality Measures. *IEEE J. Oceanic Eng.*, vol. 41, no. 3, pp. 541–551, 2016. <https://doi.org/10.1109/JOE.2015.2469915>
- [30] M. Yang, A. Sowmya. An Underwater Color Image Quality Evaluation Metric. *IEEE Trans. Image Process.*, vol. 24, no. 12, pp. 6062–6071, 2015. <https://doi.org/10.1109/TIP.2015.2491020>
- [31] R. Liu, X. Fan, M. Hou, Z. Jiang, Z. Luo, L. Zhang. Learning Aggregated Transmission Propagation Networks for Haze Removal and Beyond. *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 30, no. 10, pp. 2973–2986, 2019. <https://doi.org/10.1109/TNNLS.2018.2862631>
- [32] D. Berman, T. Treibitz, S. Avidan. Diving into Haze-Lines: Color Restoration of Underwater Images. In: *Proc. British Machine Vision Conference*, paper 44, 2017. <https://doi.org/10.5244/C.31.44>
- [33] D. Berman, T. Treibitz, S. Avidan. Non-Local Image Dehazing. In: *Proc. IEEE CVPR*, pp. 1674–1682, 2016. <https://doi.org/10.1109/CVPR.2016.185>
- [34] Q. Zhu, J. Mai, L. Shao. A Fast Single Image Haze Removal Algorithm Using Color Attenuation Prior. *IEEE Trans. Image Process.*, vol. 24, no. 11, pp. 3522–3533, 2015. <https://doi.org/10.1109/TIP.2015.2446191>