



ARTICLE

IG-Mamba: Isoline-Guided Evolutionary State Space Model for Physics-Informed Underwater Image Restoration

Yiqiao Xiang¹, Jingchun Zhou^{1,2,*}, Ruijie Liu¹ and Dehuan Zhang¹¹College of Information Science and Technology, Dalian Maritime University, Dalian, China²State Key Laboratory of Ocean Sensing, ZJU-Hangzhou Global Scientific and Technological Innovation Center, Zhejiang University, Hangzhou, China*Corresponding Author: Jingchun Zhou. Email: zhoujingchun@dlmu.edu.cn

Received: 16 March 2026; Accepted: 18 May 2026; Published: 23 July 2026

ABSTRACT: Underwater imagery is degraded by depth-dependent absorption and scattering, which often introduce color casts and contrast attenuation. Although recent Vision Mamba models provide efficient long-range dependency modeling, their conventional 2D scanning patterns are not explicitly designed to exploit the depth-correlated structure of underwater degradation and may therefore weaken geometry-aware feature dependencies. To address this limitation, we propose Isoline-Guided Evolutionary Mamba (IG-Mamba), a physics-inspired framework that uses a depth-correlated potential prior to organize state-space token propagation. Specifically, we introduce a Topology-Preserving Isoline Scanning mechanism. By leveraging a geometric prior, this mechanism quantizes the scene into discrete isopotential strata to guide Mamba sequences along geometry-aware orders, thereby preserving local spatial topology while establishing depth-aware long-range dependencies. Furthermore, a Potential Field Evolution method is developed to mitigate the domain discrepancy between terrestrial geometric priors and underwater optical attenuation. Driven by the reconstruction objective, the network adaptively refines the raw geometric anchor into a restoration-oriented optical potential field via a learned residual map. Finally, features are modulated by a transmission-inspired gate motivated by the Jaffe-McGlamery transmission term, enabling spatially adaptive feature reweighting under scattering-dominant conditions. Extensive experiments demonstrate that IG-Mamba achieves strong performance across multiple benchmarks, offering a physics-grounded perspective for dependency modeling in underwater vision.

KEYWORDS: Underwater image enhancement; state space model; isoline scanning; physics-informed learning

1 Introduction

Underwater Image Enhancement (UIE) is essential for marine operations such as autonomous underwater vehicles (AUVs) and biological observation [1–3]. According to the Jaffe-McGlamery model, underwater degradation is mainly caused by wavelength-dependent absorption and scattering, both of which are correlated with scene depth [4–6]. Although recent data-driven UIE methods achieve strong performance, many of them still organize feature interactions primarily on the 2D image plane, without explicitly exploiting the depth-conditioned structure of underwater degradation.

To examine this structure, we compute a scattering-related degradation proxy on paired UIEB images:

$$\tilde{D}_i(x) = \max(C(G_\sigma(Y_i^{\text{gt}}))(x) - C(G_\sigma(Y_i^{\text{in}}))(x), 0),$$

where Y_i^{gt} and Y_i^{in} denote the luminance representations of the reference and degraded images, respectively, G_σ denotes Gaussian low-pass filtering, and $C(\cdot)$ is a local-contrast operator implemented by local standard deviation. For aggregation, the proxy is min-max normalized within each image. We then divide pixels into five near-to-far quantile bins according to an external depth-correlated prior and compute the bin-wise mean of the normalized proxy. As shown in Fig. 1, the proxy generally increases from near to far, suggesting that underwater degradation contains depth-conditioned structure rather than only unstructured corruption.

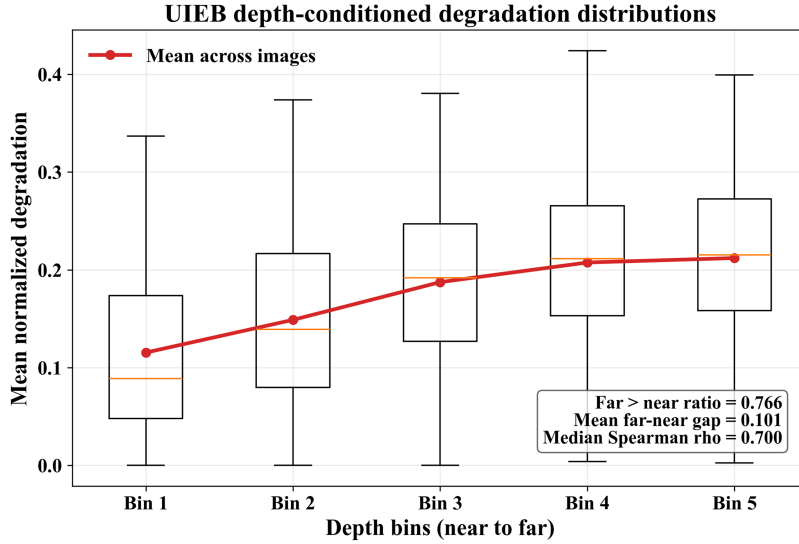


Figure 1: Depth-conditioned degradation statistics on UIEB. Pixels are divided into five near-to-far bins according to the external depth-correlated prior. The increasing trend suggests that underwater degradation exhibits depth-conditioned statistics.

Mamba [7] provides efficient long-range sequence modeling, but when applied to vision tasks, its effectiveness critically depends on how 2D tokens are organized into 1D sequences. Existing Vision Mamba variants usually adopt raster, cross, or other image-plane scan orders, which are not explicitly aligned with depth-related underwater degradation. This motivates us to move from coordinate-based scanning to geometry-aware sequence organization.

We propose IG-Mamba, a physics-inspired state-space framework for underwater image restoration. Its key component is Topology-Preserving Isoline Scanning, which partitions pixels into discrete iso-potential strata using a depth-correlated potential prior. Within each stratum, stable sorting preserves the original row-wise or column-wise local order, reducing local fragmentation while enabling state propagation along geometry-aware orders. Since terrestrial geometric priors may not match underwater optical degradation, we further introduce Potential Field Evolution to refine the coarse geometry anchor into a restoration-oriented optical potential, followed by transmission-inspired feature modulation for scattering-dominant regions.

The main contributions are summarized as follows:

- We propose IG-Mamba, which uses a depth-correlated potential field to organize the scanning topology of visual SSMs for underwater image restoration.
- We present Topology-Preserving Isoline Scanning, which guides Mamba sequences along geometry-aware potential strata and preserves local 2D order within each stratum through stable sorting.
- We introduce Potential Field Evolution and transmission-inspired feature modulation to calibrate coarse terrestrial geometry priors into restoration-oriented optical potentials.

- Extensive experiments, degraded-prior stress tests, efficiency analysis, and bin-number sensitivity studies demonstrate the effectiveness and practical feasibility of IG-Mamba.

2 Related Work

2.1 Underwater Image Enhancement

Underwater image enhancement (UIE) has been extensively studied to counteract the severe color shifts and contrast loss caused by light absorption and scattering [8]. According to the classic Jaffe-McGlamery optical model [9], underwater imaging is a depth-dependent process where the degradation is highly correlated with the distance between the camera and the reflective object. Early approaches heavily relied on hand-crafted physical priors, such as the Dark Channel Prior (DCP) [10], to estimate transmission maps. However, these prior-based methods often suffer from limited generalization due to their brittle assumptions.

In recent years, learning-based UIE methods have dominated the field [11–14]. While achieving significant improvements in visual quality, many learning-based UIE methods still formulate restoration mainly as an image-to-image mapping problem. Physical priors, when used, are often introduced as auxiliary inputs, supervision signals, or output-level constraints, rather than being deeply embedded into the feature interaction process.

2.2 Vision State Space Models

State Space Models (SSMs) derive from continuous-time systems, and Mamba [7] excels at modeling long-term dependencies. Applying 1D sequence models to 2D vision problems generally requires a spatial scanning strategy, such as the four-directional cross-scan in SS2D [15].

However, existing methods generally treat Mamba as a general-purpose feature extractor and organize 2D feature maps into 1D sequences along predetermined image-plane paths. Such coordinate-defined scanning orders do not explicitly align the state-propagation path of SSMs with the depth-dependent optical degradation in underwater scenes. Consequently, scanning feature maps across different degradation states with a fixed image-plane order may be less aligned with underwater degradation statistics.

2.3 Physics-Driven Deep Learning and Depth Priors

Recognizing the critical role of 3D physical geometry in underwater degradation, previous UIE studies have attempted to incorporate depth or physical priors through multi-task depth estimation, feature concatenation, region-level guidance, or physical constraints [16–18]. These methods demonstrate that geometry-related cues are useful for UIE, but the prior usually participates as an auxiliary input, an additional supervision signal, or a constraint on the restored image. In contrast, IG-Mamba uses a depth-correlated potential field to define the token ordering of the SSM sequence itself. Thus, the prior does not merely enrich the input representation; it organizes the state-propagation topology of Mamba through isopotential routing. This distinction shifts depth guidance from feature injection to geometry-aware sequence organization. To avoid training an underwater depth estimator from scratch, our implementation uses DepthAnythingV2 [19] as a default prior provider, while later experiments further show that the framework is not tied to this specific estimator.

3 Methodology

3.1 Overall Architecture

The overall architecture of the proposed Isoline-Guided Evolutionary Mamba (IG-Mamba) is illustrated in Fig. 2. Unlike many UIE networks that rely solely on degraded color images, IG-Mamba adopts a physics-informed dual-input strategy. Given a degraded underwater image $I_{in} \in \mathbb{R}^{B \times 3 \times H \times W}$ and its corresponding depth-correlated geometry prior $P_{in} \in \mathbb{R}^{B \times 1 \times H \times W}$, our objective is to reconstruct the enhanced image I_{out} . Specifically, I_{in} is processed by a convolutional stem to extract initial visual features $X_0 \in \mathbb{R}^{B \times C \times H \times W}$. Concurrently, P_{in} serves as the initial potential field P_0 for the routing and modulation stream. As shown in Fig. 2, the backbone follows a multi-stage encoder-decoder structure: the encoder consists of IGStage modules, while the decoder uses corresponding IGStage_up modules with skip connections from the encoder. Each stage contains stacked IGBlock units for geometry-guided restoration, followed by spatial downsampling or upsampling operations. For notation consistency, IGBlock denotes the basic unit containing IsolineSort-guided Mamba modeling, local detail preservation, potential updating, and transmission-inspired modulation.

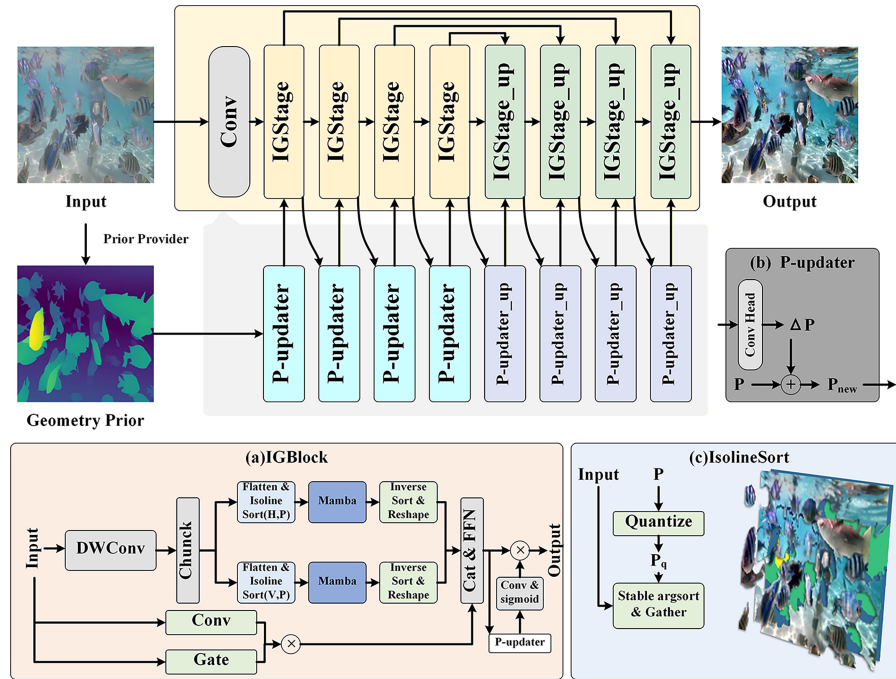


Figure 2: Overall architecture of IG-Mamba. An external geometry prior initializes the potential stream, which is propagated together with the visual feature stream through the encoder-decoder backbone. IGStage and IGStage_up denote the encoder and decoder stages. (a) IGBlock combines local convolution, horizontal and vertical IsolineSort-guided Mamba branches, gated local-detail preservation, P -updating, and transmission-inspired feature modulation. (b) The P -updater predicts ΔP and updates the potential by $P_{new} = P + \Delta P$. (c) IsolineSort quantizes P into P_q and uses stable argsort to build geometry-aware token sequences while preserving local order within each iso-potential stratum.

For clarity, P_0 , P , P_q , and P_{new} denote different computational states of the same potential stream. P_0 is the external initial geometry prior, P is the current potential field received by a block, P_q is the quantized potential used only for isoline sorting, and $P_{new} = P + \Delta P$ is the evolved potential used for subsequent propagation and physical modulation. The framework does not assume P_0 to be metrically accurate depth; it serves as a coarse spatial geometry anchor.

To maintain a stable scale space for the potential field across hierarchical levels, the potential stream employs average pooling for spatial downsampling and bilinear interpolation for upsampling. During skip connections, element-wise mean fusion is applied ($P_{\text{out}} = \frac{P_{\text{up}} + P_{\text{skip}}}{2}$) instead of standard addition to prevent value inflation beyond a stable numerical range.

3.2 Physics-Guided Isoline Routing Mamba

Optical Triangle and Radiative-Transfer-Motivated State Propagation. In underwater environments, light propagation follows the optical triangle principle: light travels from the source to the object (distance d_1), reflects, and returns to the camera (distance d_2). Based on the Jaffe-McGlamery model [9], visual deterioration in a stationary state is primarily governed by the camera-to-object distance. Recently, most Vision Mamba models have converted two-dimensional spatial information into one-dimensional sequences using fixed scanning pathways, such as rasterization. However, spatial-agnostic scanning may connect pixels with different depth-related degradation states in the SSM sequence, thereby weakening local structural integrity. To resolve this issue, we propose the *Topology-Preserving Isoline Scanning* scheme based on the one-dimensional line-of-sight Radiative Transfer Equation (RTE) theory. In underwater optics, the propagation of light at a distance z obeys the following law:

$$\frac{dL}{dz} = -c \cdot L(z) + S(z), \quad (1)$$

Here, z denotes the physical optical-path coordinate. Notably, the continuous-time SSM equation has a similar recursive form:

$$h'(t) = Ah(t) + Bx(t). \quad (2)$$

In IG-Mamba, the learned potential P is not treated as the physical optical-path coordinate. Instead, it is used as a depth-correlated ordering variable to group pixels into near-to-far iso-potential strata. By grouping pixels into near-to-far iso-potential strata, IsolineSort constructs geometry-aware token sequences in which later hidden states can aggregate information from preceding tokens with related optical paths or degradation states. This provides a physics-inspired inductive bias analogous to radiative transfer, where radiance accumulates along an optical path through attenuation and scattering.

Orthogonal Isoline Routing. To approximate this ordered accumulation over discrete tensors while better preserving local two-dimensional structure, we quantize the potential field P into discrete ordinal iso-potential indices with a bin granularity controlled by K . Specifically, $\text{Quantize}(\cdot, K)$ maps potential values into discrete ordinal indices, where pixels assigned to the same index are treated as belonging to the same iso-potential stratum. To address potential directional bias, the input features are decomposed by channel into horizontal (X_h) and vertical (X_v) branches, respectively.

The resulting operation, termed *IsolineSort*, applies stable sorting according to the quantized potential rather than exact metric depth. Stability ensures that pixels within the same iso-potential stratum retain their original row-wise and column-wise local orders in the X_h and X_v branches, respectively. These visual tokens are then flattened to form one-dimensional, isoline-ordered sequences. After independent state-space evolution by Mamba_H and Mamba_V , the sequences are mapped back to their original 2D spatial grids by inverse sorting and then concatenated. The detailed dual-route algorithm is shown in Algorithm 1. Here, $(\cdot)^T$ denotes transposition of the two spatial dimensions H and W .

Algorithm 1: Topology-preserving isoline scanning

Require: Visual Feature $\mathbf{X} \in \mathbb{R}^{B \times C \times H \times W}$, Optical Potential $\mathbf{P} \in \mathbb{R}^{B \times 1 \times H \times W}$, Bin Count K

Ensure: Routed Feature $\mathbf{Y}_{\text{route}} \in \mathbb{R}^{B \times C \times H \times W}$

- 1: **Feature Split & Potential Quantization**
 - 2: $\mathbf{X}_h, \mathbf{X}_v \leftarrow \text{Split}(\mathbf{X}, \text{dim} = C)$
 - 3: $\mathbf{P}_q \leftarrow \text{Quantize}(\mathbf{P}, K)$
 - 4: **Isoline Index Generation**
 - 5: $\Pi_h \leftarrow \text{Argsort}(\text{Flatten}(\mathbf{P}_q), \text{stable} = \text{True})$
 - 6: $\Pi_v \leftarrow \text{Argsort}(\text{Flatten}(\mathbf{P}_q^\top), \text{stable} = \text{True})$
 - 7: **Routed State Space Evolution**
 - 8: $\tilde{\mathbf{Y}}_h \leftarrow \text{Mamba}_H(\text{Gather}(\text{Flatten}(\mathbf{X}_h), \Pi_h))$
 - 9: $\tilde{\mathbf{Y}}_v \leftarrow \text{Mamba}_V(\text{Gather}(\text{Flatten}(\mathbf{X}_v^\top), \Pi_v))$
 - 10: **Inverse Mapping & Feature Fusion**
 - 11: $\Pi_h^{-1}, \Pi_v^{-1} \leftarrow \text{Argsort}(\Pi_h, \text{stable} = \text{True}), \text{Argsort}(\Pi_v, \text{stable} = \text{True})$
 - 12: $\mathbf{Y}_h \leftarrow \text{Reshape}_{H \times W}(\text{Gather}(\tilde{\mathbf{Y}}_h, \Pi_h^{-1}))$
 - 13: $\mathbf{Y}_v \leftarrow (\text{Reshape}_{W \times H}(\text{Gather}(\tilde{\mathbf{Y}}_v, \Pi_v^{-1})))^\top$
 - 14: **return** $\text{Concat}(\mathbf{Y}_h, \mathbf{Y}_v, \text{dim} = C)$
-

3.3 Parallel Local Detail Branch

While the isoline-guided SSM captures macro-level optical dependencies, it might inherently overlook localized, high-frequency textures that are orthogonal to the depth gradient (e.g., coral textures). We introduce a parallel Local Detail Branch utilizing a 3×3 depth-wise convolution and Channel Attention:

$$X_{\text{local}} = \mathcal{H}_{\text{conv}}(X) \odot \sigma(\mathcal{H}_{\text{gate}}(\text{GAP}(\mathcal{H}_{\text{conv}}(X)))), \quad (3)$$

The aggregated feature is fused via a learnable scaling parameter α : $X_{\text{fused}} = \text{FFN}(\mathbf{Y}_{\text{route}} + X_{\text{local}} \cdot \alpha) + X$. This dual-branch design supports both global geometry-aware modeling and local visual fidelity.

3.4 Potential Field Evolution and Transmission-Inspired Modulation

Although pre-trained foundation models can provide useful geometric priors, a non-negligible domain gap remains between in-air geometric depth and underwater optical degradation, which is affected by turbidity, backscattering, and wavelength-dependent absorption. The same depth-dependent degradation principle also motivates a complementary modulation path: while geometry-aware routing uses the potential to define the state-propagation order, the Jaffe-McGlamery transmission term $t(x) = e^{-\beta d(x)}$ suggests that depth-correlated geometry can also provide an attenuation-related signal for feature reweighting.

To implement this idea while reducing the geometry-optics domain gap, IG-Mamba proposes the *Potential Field Evolution* (ΔP) mechanism (Algorithm 2). During the forward pass, the coarse geometric prior is refined by a learned residual correction ΔP to obtain the updated optical potential field $P_{\text{new}} \in \mathbb{R}^{B \times 1 \times H \times W}$. Here, P_{new} should be interpreted as a restoration-oriented, degradation-aware optical potential rather than a metrically accurate depth map. A lightweight convolutional mapping then expands P_{new} into a spatial-channel transmission-inspired gate $W_{\text{mod}} \in [0, 1]^{B \times C \times H \times W}$ for feature reweighting.

The aggregated visual features are then modulated through element-wise multiplication, $X_{\text{fused}} \odot W_{\text{mod}}$. This differentiable branch also plays an optimization role beyond its physical motivation. Since the `argsort` operations in our isoline routing are non-differentiable and prevent direct gradient propagation through sorting indices, W_{mod} provides an additional differentiable optimization path. Consequently, improper modulation, such as over-suppressing a bright foreground region, increases the reconstruction loss, which backpropagates through W_{mod} to adaptively calibrate the potential field. Radiometric compensation is mainly handled by the main restoration stream and reconstruction head, while the modulation branch provides transmission-inspired reweighting for scattering-dominant feature responses and encourages the potential stream to retain its association with depth-correlated geometry.

Finally, this design forms a restoration-driven correction loop: the routing path uses the current potential to organize state propagation, while the modulation path allows the reconstruction objective to adapt the potential through a continuous feature-reweighting branch. This encourages ΔP to adjust the raw geometric prior toward a restoration-oriented optical potential without requiring auxiliary depth supervision.

Algorithm 2: Transmission-inspired feature reweighting

Require: Fused Feature $\mathbf{X}_{\text{fused}} \in \mathbb{R}^{B \times C \times H \times W}$, Optical Potential $\mathbf{P} \in \mathbb{R}^{B \times 1 \times H \times W}$
Ensure: Final Feature $\mathbf{X}_{\text{final}} \in \mathbb{R}^{B \times C \times H \times W}$, Evolved Potential $\mathbf{P}_{\text{new}} \in \mathbb{R}^{B \times 1 \times H \times W}$

- 1: **Potential Field Evolution via Semantic Residuals**
- 2: $\Delta \mathbf{P} \leftarrow \text{Conv}_{3 \times 3}(\text{GELU}(\text{Conv}_{1 \times 1}(\mathbf{X}_{\text{fused}})))$
- 3: $\mathbf{P}_{\text{new}} \leftarrow \mathbf{P} + \Delta \mathbf{P}$ {*Fine-grained optical calibration*}
- 4: **2. Transmission-Inspired Gate Generation**
- 5: $\mathbf{W}_{\text{mod}} \leftarrow \sigma(\text{Conv}_{3 \times 3}^{1 \rightarrow C}(\mathbf{P}_{\text{new}}))$ { $\mathbf{W}_{\text{mod}} \in (0, 1)^{B \times C \times H \times W}$ }
- 6: **3. Transmission-Inspired Feature Reweighting**
- 7: $\mathbf{X}_{\text{final}} \leftarrow \mathbf{X}_{\text{fused}} \odot \mathbf{W}_{\text{mod}}$ {*transmission-inspired feature modulation*}
- 8: **return** $\mathbf{X}_{\text{final}}, \mathbf{P}_{\text{new}}$

3.5 Supervision Strategy of Potential Field

We train IG-Mamba end-to-end with a reconstruction-driven potential evolution strategy. The external prior P_0 generated by DepthAnythingV2 [19] is used only to initialize the potential stream as a coarse geometry anchor. Through IsolineSort-guided routing and the differentiable transmission-inspired modulation branch, the image restoration objective optimizes the potential stream and encourages ΔP to adapt the raw geometric prior into a restoration-oriented optical potential that is more consistent with underwater optical degradation.

The overall loss is entirely specified as the L_1 reconstruction loss of the image:

$$\mathcal{L}_{\text{total}} = \mathcal{L}_1 = \frac{1}{N} \|I_{\text{pred}} - I_{\text{gt}}\|_1 \quad (4)$$

where N is the total number of pixels in I_{pred} .

This depth-supervision-free design allows the potential field to be optimized for image restoration rather than being forced to fit potentially inaccurate metric depth labels.

4 Experiments

4.1 Experimental Setup

Datasets. Training and in-domain validation are conducted on UIEB [11]: 800 paired images are used for training (UIEB-T800), 90 paired images for full-reference validation (UIEB-V90), and 60 challenging images

without references for UIEB-C60 evaluation. Cross-domain non-reference evaluation is further conducted on U45 [20], UCCS [21], MABLs [22], UFO120 [23], and Z700 [24], covering different color casts, turbidity levels, and real underwater scenes.

Implementation Details. IG-Mamba is implemented in PyTorch and trained end-to-end on an NVIDIA vGPU (32 GB). We use Adam with $\beta_1 = 0.9$, $\beta_2 = 0.999$, an initial learning rate of 1×10^{-4} , and cosine annealing to 1×10^{-6} . The network is trained for 600 epochs with a batch size of 4. Images are randomly cropped to 256×256 and augmented by random horizontal/vertical flips. The external prior is provided as an input guidance field, and the model is optimized by image reconstruction loss.

The default external prior is generated by DepthAnythingV2 and precomputed before restoration experiments. It is normalized as a depth-correlated potential input and is not used as a supervised depth label. In the degraded-prior robustness study, we keep the trained IG-Mamba checkpoint fixed and only replace or corrupt P_0 at inference time, so the results evaluate test-time tolerance to imperfect prior providers rather than robustness obtained from prior augmentation.

For competing methods, we use official results or released models when available; otherwise, reproduced outputs are evaluated with the same metric scripts and testing splits. This keeps the full-reference and non-reference comparisons under a consistent protocol.

Evaluation Metrics. PSNR [25] and SSIM [26] are used for full-reference evaluation, and RMSE is additionally reported on UIEB-V90. UCIQE [27] and URanker [28] are used for non-reference evaluation. Higher values indicate better performance for PSNR, SSIM, UCIQE, and URanker, while lower values indicate better performance for RMSE.

4.2 Comparison with State-of-the-Art Methods

We compare IG-Mamba with classic prior-/CNN-based methods (WaterNet [11], UWCNN [12], Ucolor [13], SPDF [29]), Transformer/advanced restoration methods (U-shape [14], AST [30], VQC-NIR [31]), and recent competitive models including FMambaIR [32], UDNet [33], and CDF-UIE [34].

Quantitative Evaluation. Tables 1 and 2 report the quantitative comparison results. Benefiting from geometry-aware routing and global state-space modeling, IG-Mamba achieves consistently strong performance across the evaluated datasets. Compared with SSM-based methods that rely on generic image-plane scanning paths, such as FMambaIR [32], IG-Mamba obtains higher PSNR on UIEB-V90, suggesting that organizing tokens by a depth-correlated potential field provides a useful inductive bias for depth-dependent underwater degradation.

In addition to strong restoration performance, the core enhancement network of IG-Mamba maintains an efficient structure, with **5.32M parameters** and **13.02G FLOPs** at a resolution of 256×256 . The default external prior in our experiments is precomputed, and the framework is not tied to one specific depth estimator. In practice, the coarse geometric anchor may be provided by a lightweight learned proxy, low-resolution prior, or on-board AUV sensing sources such as stereo cameras and sonars. This deployment flexibility is further analyzed in Section 4.4.

Table 1: Quantitative comparison of different methods on UIEB-V90 and UIEB-C60 datasets. The 1st, 2nd, and 3rd best results are highlighted in **red**, **blue**, and **green** backgrounds, respectively. Ties are assigned the same rank.

Method	Venue	UIEB-V90					UIEB-C60	
		PSNR↑	SSIM↑	RMSE↓	URanker↑	UCIQE↑	URanker↑	UCIQE↑
WaterNet [11]	TIP'19	17.349	0.813	0.144	1.471	0.606	0.931	0.591
UWCNN [12]	PR'20	17.985	0.844	0.134	1.309	0.554	0.499	0.530
Ucolor [13]	TIP'21	20.962	0.864	0.097	1.534	0.591	0.841	0.565
SPDF [29]	TCSVT'23	20.331	0.828	0.103	1.433	0.570	1.034	0.569
U-shape [14]	TIP'23	20.462	0.792	0.100	1.665	0.576	1.059	0.553
VQCNIR [31]	AAAI'24	23.330	0.914	0.074	2.069	0.615	1.446	0.588
AST [30]	CVPR'24	21.695	0.869	0.089	1.791	0.586	1.042	0.558
FMambaIR [32]	TGRS'25	22.305	0.865	0.083	2.096	0.613	1.489	0.588
UDNet [33]	ESWA'25	19.017	0.768	0.118	1.216	0.585	0.618	0.568
CDF-UIE [34]	TGRS'25	21.699	0.808	0.088	2.042	0.603	1.387	0.576
IG-Mamba	Ours	24.521	0.907	0.069	2.210	0.621	1.573	0.591

Table 2: Quantitative comparison of different methods on U45, Z700, MABLs, UFO120, and UCCS datasets. The 1st, 2nd, and 3rd best results are highlighted in **red**, **blue**, and **green** backgrounds, respectively. Ties are assigned the same rank.

Method	Venue	U45		Z700		MABLs		UFO120		UCCS	
		URanker↑	UCIQE↑	URanker↑	UCIQE↑	URanker↑	UCIQE↑	URanker↑	UCIQE↑	URanker↑	UCIQE↑
WaterNet [11]	TIP'19	1.330	0.594	1.707	0.559	1.448	0.600	1.748	0.615	1.216	0.565
UWCNN [12]	PR'20	1.513	0.554	1.495	0.503	1.030	0.532	1.893	0.579	0.983	0.501
Ucolor [13]	TIP'21	1.471	0.586	1.895	0.543	1.255	0.573	1.754	0.612	1.090	0.553
SPDF [29]	TCSVT'23	1.284	0.565	1.492	0.527	1.413	0.560	1.530	0.572	1.212	0.508
U-shape [14]	TIP'23	1.502	0.572	1.976	0.553	1.514	0.561	1.678	0.592	1.178	0.528
VQCNIR [31]	AAAI'24	2.067	0.616	2.106	0.577	1.812	0.592	2.122	0.635	1.371	0.550
AST [30]	CVPR'24	1.776	0.589	1.807	0.544	1.363	0.564	1.938	0.608	0.827	0.523
FMambaIR [32]	TGRS'25	1.853	0.609	1.937	0.574	1.811	0.590	2.175	0.636	1.267	0.548
UDNet [33]	ESWA'25	0.858	0.584	1.299	0.558	0.997	0.569	1.506	0.602	0.765	0.555
CDF-UIE [34]	TGRS'25	1.915	0.603	2.091	0.563	1.798	0.583	2.134	0.624	1.306	0.547
IG-Mamba	Ours	2.029	0.618	2.292	0.593	1.935	0.599	2.134	0.635	1.380	0.554

Qualitative Evaluation. Figs. 3–5 provide visual comparisons on UIEB-V90 and cross-domain datasets. IG-Mamba reduces color casts and haze-like scattering while preserving foreground structures and object details, showing robust restoration under diverse underwater conditions.

4.3 Robustness to Degraded External Geometry Priors

To address the concern that IG-Mamba may rely on an accurate external depth estimator, we conduct a controlled robustness study on UIEB-V90 by deliberately degrading the input geometry prior P_0 . Downsampled and blurred priors preserve coarse but spatially aligned geometry; noisy and masked priors simulate local estimation errors or missing regions; a constant prior represents the no-geometry fallback; shuffled and

random priors represent spatially inconsistent or fully corrupted guidance. Table 3 reports the robustness results under these degraded-prior settings.

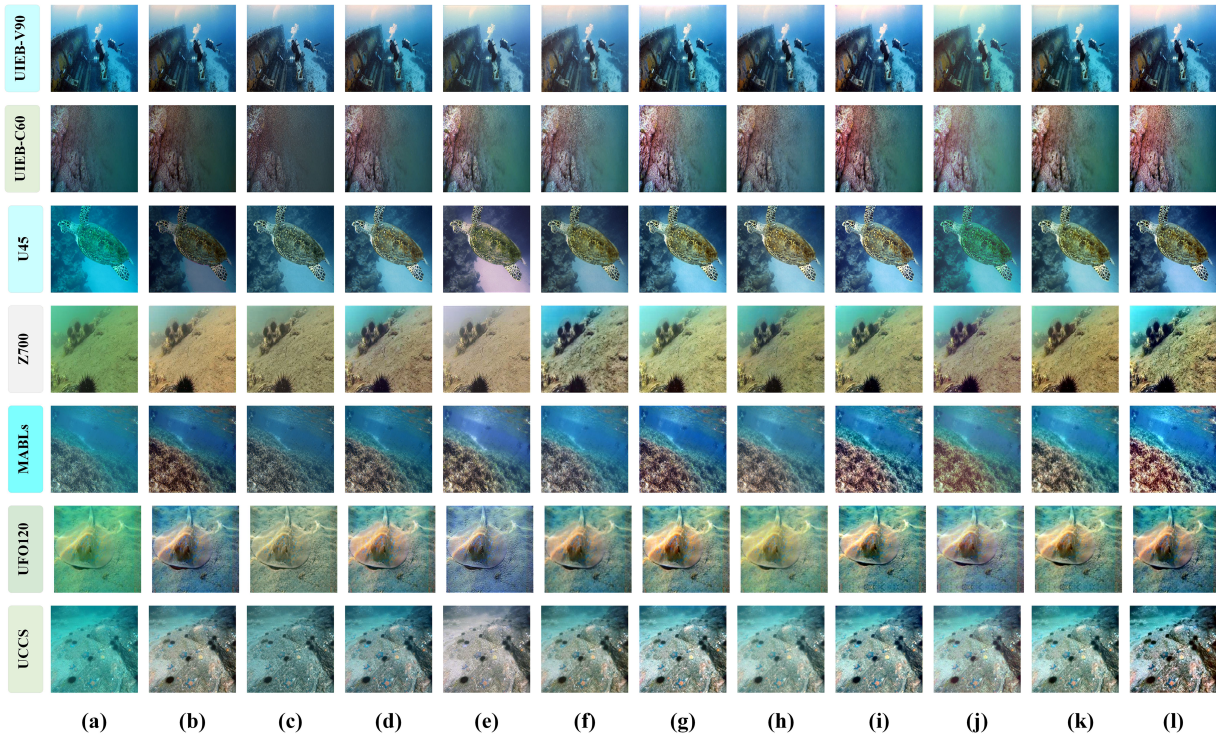


Figure 3: Qualitative comparison on reference-based and non-reference underwater image enhancement benchmarks, including UIEB-V90, UIEB-C60, U45, Z700, MABLs, UFO120, and UCCS. Columns (a–l) correspond to the input, WaterNet, UWCNN, Ucolor, SPDF, U-shape, VQCNIR, AST, FMambaIR, UDNNet, CDF-UIE, and our IG-Mamba, respectively. IG-Mamba produces visually balanced restoration results across different underwater scenes and degradation types.

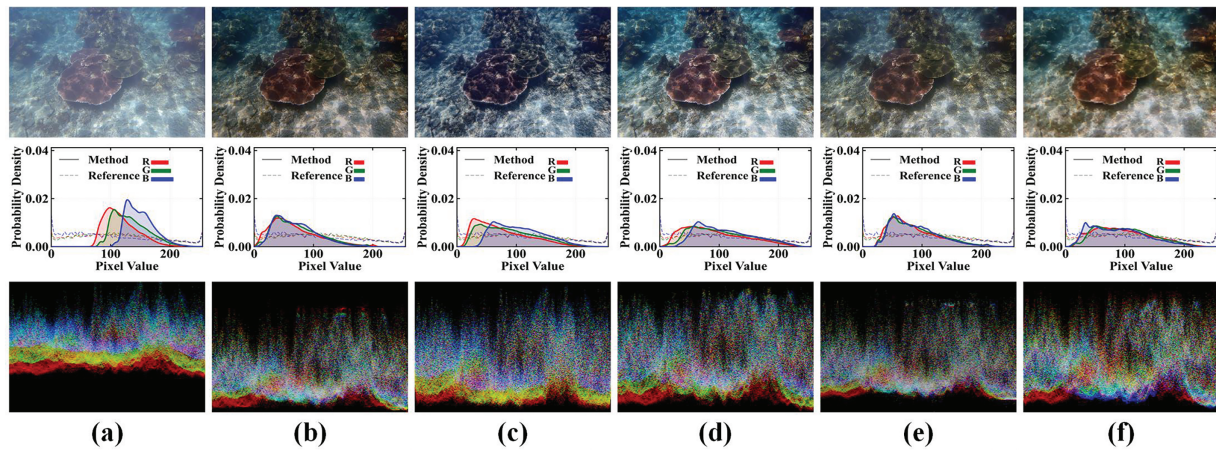


Figure 4: (Continued)

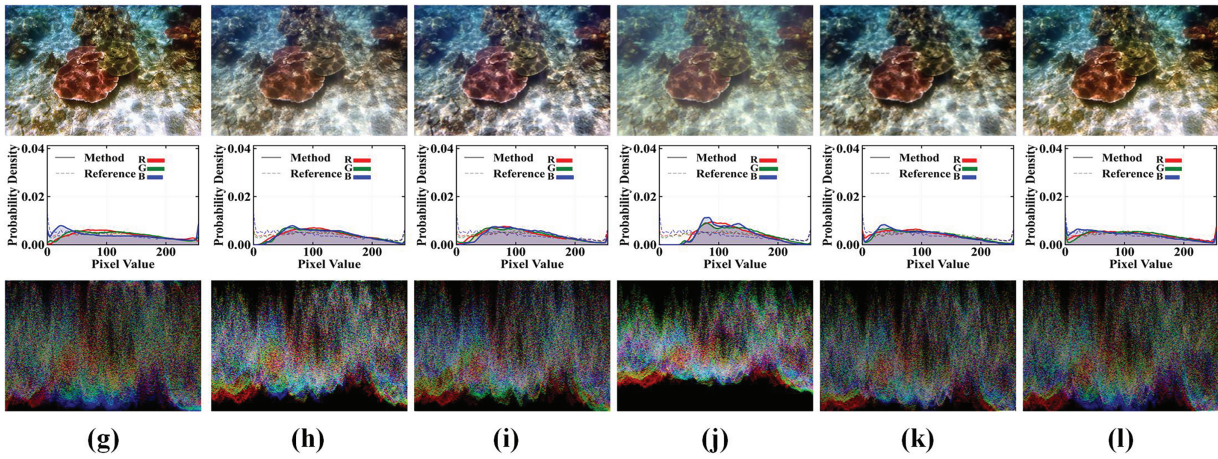


Figure 4: Visual and color-distribution comparison on a representative underwater image. Each method is shown together with its RGB probability-density curves and color distribution. Columns (a–l) correspond to the input, WaterNet, UWCNN, Ucolor, SPDF, U-shape, VQCNIR, AST, FMambaIR, UDNet, CDF-UIE, and our IG-Mamba, respectively. Compared with existing methods, IG-Mamba better aligns the restored color distribution with the reference while reducing underwater color casts.

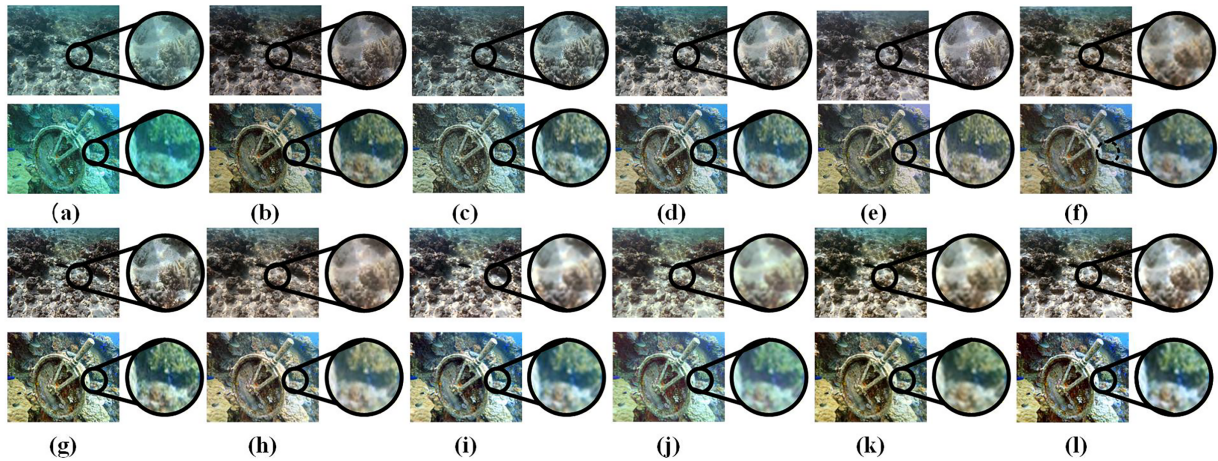


Figure 5: Qualitative comparison with local magnified regions on challenging non-reference underwater images. Columns (a–l) correspond to the input, WaterNet, UWCNN, Ucolor, SPDF, U-shape, VQCNIR, AST, FMambaIR, UDNet, CDF-UIE, and our IG-Mamba, respectively. IG-Mamba preserves clearer local structures and improves visibility under challenging underwater degradation.

The results show that IG-Mamba does not require metrically accurate depth. Moderately degraded but spatially aligned priors, such as downsampled, blurred, noisy, and masked priors, do not cause performance collapse. This suggests that a coarse relative geometry anchor is sufficient for stable iso-potential grouping. In particular, coarse or smoothed priors can remain effective because they preserve the near-to-far spatial layout while reducing high-frequency artifacts in monocular depth estimates. When the spatial layout of the prior is removed or disrupted, as in constant, shuffled, and random priors, the performance drops more noticeably, suggesting that IG-Mamba relies on spatially meaningful geometry anchors rather than metrically precise depth. Together with the subsequent Static Prior ablation study, these results indicate that potential evolution helps calibrate imperfect geometry priors into restoration-oriented optical potentials.

Table 3: Robustness of IG-Mamba to degraded external geometry priors on UIEB-V90. Each cell reports PSNR/SSIM. Δ PSNR is computed relative to the clean-prior setting. Aligned perturbations preserve the spatial layout of the prior, while non-geometric priors destroy or remove the geometry guidance.

Prior Setting	Prior Property	IG-Mamba	Δ PSNR
Clean prior	Original geometry prior	24.52/0.9066	0.00
Downsampled prior	Coarse but aligned	24.53/0.9064	+0.01
Blurred prior	Smooth but aligned	24.52/0.9065	-0.01
Noisy prior	Noisy but aligned	24.49/0.9058	-0.03
Masked prior	Partially missing but aligned	24.51/0.9058	-0.01
Constant map	No geometry information	24.14/0.9024	-0.38
Shuffled map	Spatially misaligned geometry	24.17/0.9011	-0.36
Random map	Fully corrupted guidance	24.08/0.8996	-0.44

4.4 Efficiency and Sensitivity Analysis

Table 4 reports the efficiency and prior-provider comparison. To address practical deployment concerns, we further evaluate the computational cost of different prior providers at 256×256 resolution. The external clean prior achieves the best restoration quality, but its online prior-generation cost is much higher when the DepthAnythingV2 inference time is included. In contrast, the lightweight learned prior introduces only 0.017M parameters and 0.193G FLOPs, reducing the total latency from 180.06 to 35.26 ms while maintaining competitive restoration quality. This result indicates that IG-Mamba is not tied to a heavy monocular depth foundation model; a fast task-oriented prior proxy can also provide an effective coarse geometry anchor for deployment.

Table 4: Efficiency and prior-provider analysis on UIEB-V90 at 256×256 resolution. Runtime is measured with batch size 1 after warm-up. For external priors, the measured DepthAnythingV2 prior-generation time is included. The restoration core is shared by all settings and contains 5.32M parameters and 13.02G FLOPs.

Prior Provider	Prior Params	Prior FLOPs	Prior Time	Total Time/FPS	PSNR/SSIM
External clean prior	335.316M	154.807G	145.54 ms	180.06 ms/5.55	24.521/0.9066
Lightweight learned prior	0.017M	0.193G	2.13 ms	35.26 ms/28.36	24.316/0.9036
Constant fallback	–	–	0.00 ms	33.32 ms/30.01	24.143/0.9024

We also study the sensitivity to the number of iso-potential bins K , as reported in Table 5. The sensitivity analysis shows that the no-stratification reference ($K = 1$) is clearly inferior to valid iso-potential grouping, confirming the necessity of geometry-aware sequence organization. The best performance is obtained at $K = 4$, suggesting that underwater restoration benefits from coarse but effective iso-potential strata. When K becomes larger, PSNR decreases, likely because overly fine quantization fragments spatially coherent degradation regions and amplifies high-frequency artifacts in monocular priors. This result is consistent with the degraded-prior robustness analysis: IG-Mamba requires spatially meaningful geometry guidance, but does not rely on metrically precise depth.

Table 5: Sensitivity analysis of the number of iso-potential bins K on UIEB-V90. The $K = 1$ row corresponds to the no-stratification reference from the ablation study, while the other settings are trained with the corresponding iso-potential bin count. Bold indicates the best result.

K	1	4	16	64
PSNR $\uparrow$	23.552	24.521	24.206	24.164
SSIM $\uparrow$	0.902	0.907	0.903	0.905

4.5 Ablation Study

We conduct ablation studies on UIEB to validate the main components of IG-Mamba. Table 6 analyzes routing, stable sorting, potential evolution, and modulation through a series of variants from the depth-free baseline to the full model. Here, **Model A (Baseline)** disables all geometry-related components and achieves 23.552 dB PSNR, providing the reference for evaluating the remaining variants.

Table 6: Ablation study on UIEB. Evo. denotes potential field evolution (ΔP), and Mod. denotes transmission-inspired feature modulation. Models A–F progressively examine the depth-free baseline, continuous sorting, unstable sorting, routing-only design, static-prior modulation, and full IG-Mamba. Bold indicates the best result.

Model Variant	Geometry Prior	Isoline Bins	Stable Sort	ΔP Evo.	Trans. Mod.	PSNR $\uparrow$	SSIM $\uparrow$
Model A (Baseline)	×	×	×	×	×	23.552	0.902
Model B (Continuous Potential)	✓	×	✓	✓	✓	23.957	0.898
Model C (Unstable Sort)	✓	✓	×	✓	✓	23.991	0.904
Model D (Routing Only)	✓	✓	✓	×	×	23.899	0.905
Model E (Static Prior)	✓	✓	✓	×	✓	23.682	0.903
Model F (IG-Mamba)	✓	✓	✓	✓	✓	24.521	0.907

1. Effectiveness of Topology-Preserving Isoline Routing. To test the routing algorithm, we separately examine the effects of hard-quantized iso-potential bins and stable sorting.

First, **Model B (Continuous Potential)** removes the potential quantization process and orders continuous floating-point potentials directly. Although **Model B** improves PSNR over the baseline from 23.552 to 23.957 dB, its SSIM decreases from 0.902 to 0.898. Since there are usually no identical potential values, the resulting sequence can frequently switch across distant spatial locations. This *micro-fragmentation* weakens local structure; therefore, discrete iso-potential layers are useful for preserving macroscopically continuous structures.

Second, **Model C (Unstable Sort)** uses isoline bins but employs a non-stable sorting method. Compared with the full model, PSNR decreases from 24.521 to 23.991 dB, and SSIM decreases from 0.907 to 0.904. Without stable sorting to preserve the local order within iso-potential strata, the resulting token sequence becomes more fragmented, forcing the network to compensate for more complex spatial discontinuities during reconstruction, thereby increasing the difficulty of image enhancement.

2. Synergy of Potential Evolution and Physical Modulation. We introduce two coupled physics-inspired mechanisms in our architecture: Potential Field Evolution (ΔP) for potential calibration and Transmission-Inspired Modulation (W_{mod}) for attenuation-aware feature reweighting and gradient propagation, as illustrated in Fig. 6. To fully separate their responsibilities, we examine a series of progressive variants.

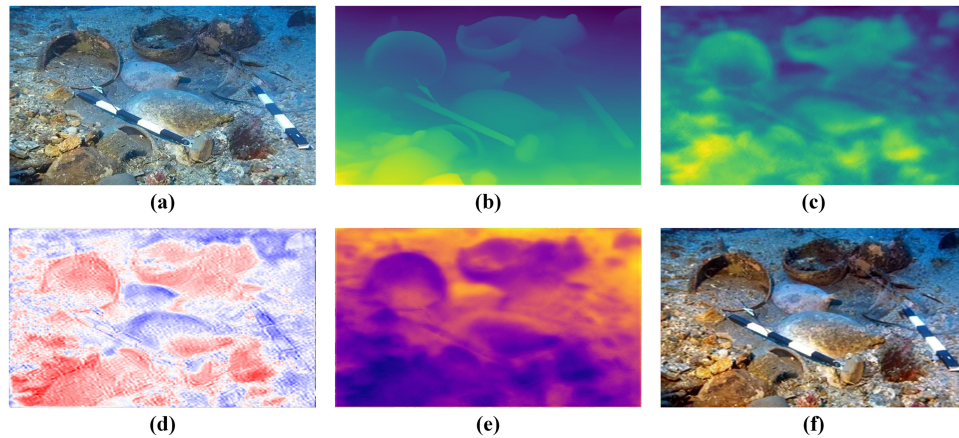


Figure 6: Visualization of the internal potential stream and modulation mechanism. From (a)–(f): input image, geometric prior P_0 , evolved optical potential P_{new} , evolution residual ΔP , modulation gate W_{mod} , and restored image. The residual ΔP calibrates P_0 into P_{new} , while W_{mod} performs transmission-inspired spatially adaptive feature reweighting for scattering-dominant regions. In the ΔP map, red and blue denote positive and negative corrections, respectively; brighter values in the W_{mod} map indicate larger gating responses.

First, **Model D (Routing Only)** uses only a static geometric prior (P_0) for Mamba sequencing. Using only the routing path, its PSNR reaches 23.899 dB, outperforming Model A. However, directly modulating features using this raw prior, as in **Model E (Static Prior)**, reduces PSNR to 23.682 dB. This *negative transfer* suggests that uncalibrated geometric distance should not be directly treated as underwater optical attenuation. Geometric depth and underwater optical degradation are correlated but not identical, and direct modulation may suppress useful foreground responses. Finally, **Model F (Full Model)** achieves the highest PSNR of 24.521 dB. This indicates that dynamic potential calibration is needed to reduce the domain gap of Model E. The W_{mod} gate serves two functions: it provides transmission-inspired spatial reweighting and establishes a continuous, differentiable optimization path around the non-differentiable `argsort` operation.

3. Representation Capacity under Deep Physical Constraints. A concern with potential-guided modulation is that explicit spatial constraints may over-restrict deep feature representation. We therefore monitor the gate statistics of W_{mod} across hierarchical levels after training. If the modulation were harmful, the network could weaken this branch by driving the gate toward an identity-like response, i.e., $W_{\text{mod}} \approx 1$. Empirically, however, W_{mod} does not collapse to such an identity-like gate. Its mean gate values range from 0.428 to 0.521 across hierarchical levels, indicating that the modulation branch remains effective from shallow encoders to deep decoders and that underwater degradation remains relevant across multiple semantic scales rather than only at low-level layers.

This observation supports the division of roles in IG-Mamba. The routing branch organizes long-range dependencies along geometry-aware strata, the potential updater adapts the external prior to underwater restoration, and the modulation branch suppresses scattering-dominant feature responses without replacing the main restoration stream. Therefore, the performance drop of the Static Prior variant mainly arises from using an uncalibrated prior for modulation, rather than from insufficient representation capacity under physical guidance.

5 Conclusion

This paper presents IG-Mamba, a physics-inspired framework that organizes Mamba state propagation with a depth-correlated potential prior. Topology-Preserving Isoline Scanning constructs geometry-aware sequences while preserving local row- and column-wise order within iso-potential strata. Potential Field Evolution further adapts the external geometry prior into a restoration-oriented optical potential for transmission-inspired feature reweighting. Experiments, degraded-prior stress tests, efficiency analysis, and bin-number sensitivity studies demonstrate that IG-Mamba provides an effective and practical geometry-aware SSM framework for underwater image restoration.

Acknowledgement: Not applicable.

Funding Statement: This work was supported in part by the National Natural Science Foundation of China (No. 62301105), in part by the Open Project of the State Key Laboratory of Ocean Sensing (No. OSKF-2025M04), in part by the Liaoning Provincial Science and Technology Plan Joint Program (Natural Science Foundation-General Program) (No. 2025-MSLH-113), in part by the Fundamental Research Funds for the Central Universities (No. 3132026240 and 3132025268), in part by the Fundamental Research Funds of the Liaoning Provincial Department of Education (No. LJ212510151021).

Author Contributions: The authors confirm contribution to the paper as follows: conceptualization, methodology, software, formal analysis, visualization, and writing—original draft: Yiqiao Xiang; data curation and validation: Yiqiao Xiang, Ruijie Liu and Dehuan Zhang; writing—review and editing: Jingchun Zhou, Ruijie Liu and Dehuan Zhang; supervision, project administration, and funding acquisition: Jingchun Zhou. All authors reviewed and approved the final version of the manuscript.

Availability of Data and Materials: The authors confirm that the data supporting the findings of this study are available within the article. The public benchmark datasets used in this study are available from their original sources cited in the article.

Ethics Approval: Not applicable. This study did not involve human participants or animals.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Song W, Wang H, Zhang X, Xia J, Liu T, Shi Y. Deep-sea nodule mineral image segmentation algorithm based on Pix2PixHD. *Comput Mater Contin.* 2022;73(1):1449–62. doi:10.32604/cmc.2022.027213.
2. Li C, Wang J, Ding C, Ye Z. AquaTree: deep reinforcement learning-driven Monte Carlo tree search for underwater image enhancement. *Comput Mater Contin.* 2026;86(3):61.
3. Zhou J, Zhang D, He Z, Gai Q, Jiang Q. Multi-prior fusion transfer plugin for adapting in-air models to underwater image enhancement and detection. *IEEE Trans Image Process.* 2025;34:7773–85. doi:10.1109/tip.2025.3634001.
4. Yang T, Jia S, Ma H. Research on the application of super resolution reconstruction algorithm for underwater image. *Comput Mater Contin.* 2020;62(3):1249. doi:10.32604/cmc.2020.05777.
5. Zhu X, Zhang J, Zeng H. FENet: underwater image enhancement via frequency domain enhancement and edge-guided refinement. *Comput Mater Contin.* 2026;86(2):1.
6. Zhou J, He Z, Zhang D, Liu S, Fu X, Li X. Spatial residual for underwater object detection. *IEEE Trans Pattern Anal Mach Intell.* 2025;47(6):4996–5013. doi:10.1109/tpami.2025.3548652.
7. Gu A, Dao T. Mamba: linear-time sequence modeling with selective state spaces. In: *Proceedings of the First Conference on Language Modeling*; 2024 Oct 7–10; Philadelphia, PA, USA.
8. Akkaynak D, Treibitz T. Sea-thru: a method for removing water from underwater images. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*; 2019 Jun 15–20; Long Beach, CA, USA. p. 1682–91.

9. Jaffe JS. Computer modeling and the design of optimal underwater imaging systems. *IEEE J Ocean Eng.* 1990;15(2):101–11. doi:10.1109/48.50695.
10. He K, Sun J, Tang X. Single image haze removal using dark channel prior. *IEEE Trans Pattern Anal Mach Intell.* 2010;33(12):2341–53. doi:10.1109/tpami.2010.168.
11. Li C, Guo C, Ren W, Cong R, Hou J, Kwong S, et al. An underwater image enhancement benchmark dataset and beyond. *IEEE Trans Image Process.* 2019;29:4376–89. doi:10.1109/TIP.2019.2955241.
12. Li C, Anwar S, Porikli F. Underwater scene prior inspired deep underwater image and video enhancement. *Pattern Recognit.* 2020;98(1):107038. doi:10.1016/j.patcog.2019.107038.
13. Li C, Anwar S, Hou J, Cong R, Guo C, Ren W. Underwater image enhancement via medium transmission-guided multi-color space embedding. *IEEE Trans Image Process.* 2021;30:4985–5000. doi:10.1109/tip.2021.3076367.
14. Peng L, Zhu C, Bian L. U-shape transformer for underwater image enhancement. *IEEE Trans Image Process.* 2023;32(2):3066–79. doi:10.1109/tip.2023.3276332.
15. Liu Y, Tian Y, Zhao Y, Yu H, Xie L, Wang Y, et al. Vmamba: visual state space model. *Adv Neural Inf Process Syst.* 2024;37:103031–63.
16. Yang G, Kang G, Lee J, Cho Y. Joint-ID: transformer-based joint image enhancement and depth estimation for underwater environments. *IEEE Sens J.* 2023;24(3):3113–22.
17. Du D, Si L, Xu F, Niu J, Sun F. A physical model-guided framework for underwater image enhancement and depth estimation. *arXiv:2407.04230.* 2024.
18. Huang Z, Wang X, Xu C, Li J, Feng L. Underwater variable zoom: depth-guided perception network for underwater image enhancement. *Expert Syst Appl.* 2025;259:125350.
19. Yang L, Kang B, Huang Z, Zhao Z, Xu X, Feng J, et al. Depth anything v2. *Adv Neural Inf Process Syst.* 2024;37:21875–911. doi:10.52202/079017-0688.
20. Li H, Li J, Wang W. A fusion adversarial underwater image enhancement network with a public test dataset. *arXiv:1906.06819.* 2019.
21. Liu R, Fan X, Zhu M, Hou M, Luo Z. Real-world underwater enhancement: challenges, benchmarks, and solutions under natural light. *IEEE Trans Circuits Syst Video Technol.* 2020;30(12):4861–75.
22. Song W, Wang Y, Huang D, Liotta A, Perra C. Enhancement of underwater images with statistical model of background light and optimization of transmission map. *IEEE Trans Broadcast.* 2020;66(1):153–69. doi:10.1109/tbc.2019.2960942.
23. Islam MJ, Luo P, Sattar J. Simultaneous enhancement and super-resolution of underwater imagery for improved visual perception. *arXiv:2002.01155.* 2020.
24. Zhou J, Sun J, Li C, Jiang Q, Zhou M, Lam KM, et al. HCLR-Net: hybrid contrastive learning regularization with locally randomized perturbation for underwater image enhancement. *Int J Comput Vis.* 2024;132(10):4132–56.
25. Korhonen J, You J. Peak signal-to-noise ratio revisited: is simple beautiful?. In: 2012 Fourth International Workshop on Quality of Multimedia Experience; 2012 Jul 5–7; Melbourne, VIC, Australia. p. 37–8.
26. Wang Z, Bovik AC, Sheikh HR, Simoncelli EP. Image quality assessment: from error visibility to structural similarity. *IEEE Trans Image Process.* 2004;13(4):600–12. doi:10.1109/tip.2003.819861.
27. Yang M, Sowmya A. An underwater color image quality evaluation metric. *IEEE Trans Image Process.* 2015;24(12):6062–71. doi:10.1109/tip.2015.2491020.
28. Guo C, Wu R, Jin X, Han L, Zhang W, Chai Z, et al. Underwater ranker: learn which is better and how to be better. *Proc AAAI Conf Artif Intell.* 2023;37:702–9.
29. Kang Y, Jiang Q, Li C, Ren W, Liu H, Wang P. A perception-aware decomposition and fusion framework for underwater image enhancement. *IEEE Trans Circuits Syst Video Technol.* 2022;33(3):988–1002. doi:10.1109/tcsvt.2022.3208100.
30. Zhou S, Chen D, Pan J, Shi J, Yang J. Adapt or perish: adaptive sparse transformer with attentive feature refinement for image restoration. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR); 2024 Jun 16–22; Seattle, WA, USA.* p. 2952–63.
31. Zou W, Gao H, Ye T, Chen L, Yang W, Huang S, et al. VQCNIR: clearer night image restoration with vector-quantized codebook. *Proc AAAI Conf Artif Intell.* 2024;38(7):7873–81. doi:10.1609/aaai.v38i7.28623.

32. Luan X, Fan H, Wang Q, Yang N, Liu S, Li X, et al. FMambaIR: a hybrid state space model and frequency domain for image restoration. *IEEE Trans Geosci Remote Sens.* 2025;63:4201614.
33. Saleh A, Sheaves M, Jerry D, Azghadi MR. Adaptive deep learning framework for robust unsupervised underwater image enhancement. *Expert Syst Appl.* 2025;268(1):126314. doi:10.1016/j.eswa.2024.126314.
34. Zhang H, Xu H, Yu X, Zhang X, Gao X, Wu C. CDF-UIE: leveraging cross-domain fusion for underwater image enhancement. *IEEE Trans Geosci Remote Sens.* 2025;63:4203715. doi:10.1109/tgrs.2025.3553557.