D2Mamba: Dual Domain Guided Informed Search in State Space Model for Underwater Image Enhancement

Alik Pramanick, Soumajit Roy, Arijit Sur
Indian Institute of Technology Guwahati

WACV 2026 TUCSON, AZ 3/6 – 3/10

Motivation

- Underwater images suffer from color distortion, haziness, and low-contrast due to light absorption and scattering.
- Despite deep learning advances in enhancement, challenges persist in efficiency, global context modeling, spatial-spectral consistency, and perceptually accurate detail recovery.

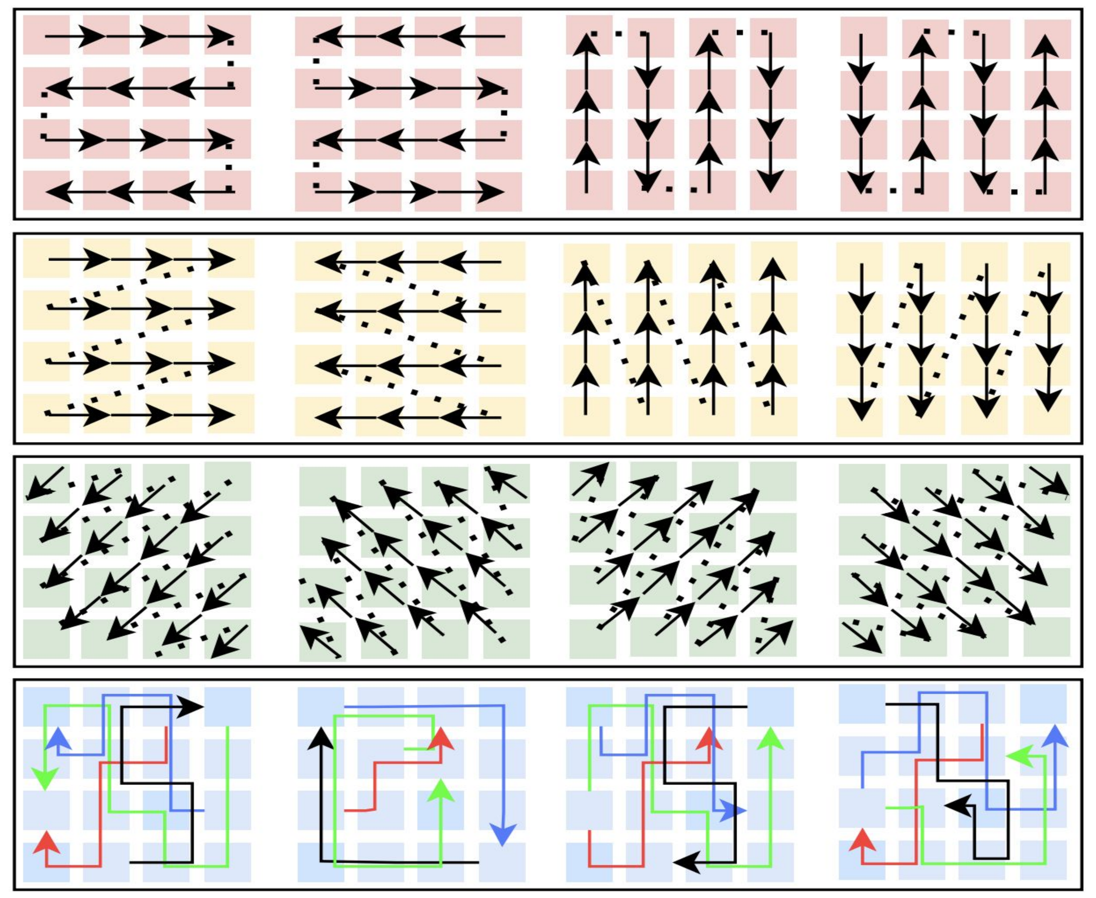

Figure 1: Comparison of different scanning strategies for MAMBA BLOCK. Unlike (a) bidirectional, (b) raster, or (c) diagonal, our (d) A* guided GIFH traversal adaptively follows degradation-aware paths, enabling more effective and sparse feature propagation.

Contributions

- Design a novel underwater image enhancement framework, D2Mamba, adopting a dual-domain information (spatial and frequency) with state space models (SSMs), enabling efficient global context modeling while preserving local de- tails.
- Unlike conventional SSMs that rely on raster, bidirectional, cross or diagonal scans, D2Mamba uses an A* search guided by physics-based Geodesic Information- Field Heuristic (GIFH) scan for feature traversal based on input degradation characteristics
- GIFH combines feature gradients, high-frequency heterogeneity, and low-frequency semantic distance to compute adaptive costs, enabling the capture of both spatial and spectral dependencies
- Further, a Spectral Wasserstein Attenuation Loss (SWAL) is introduced to enforce distributional alignment in the spectral domain, enabling perceptually consistent and physically consistent color restoration in enhanced underwater.

Proposed Method

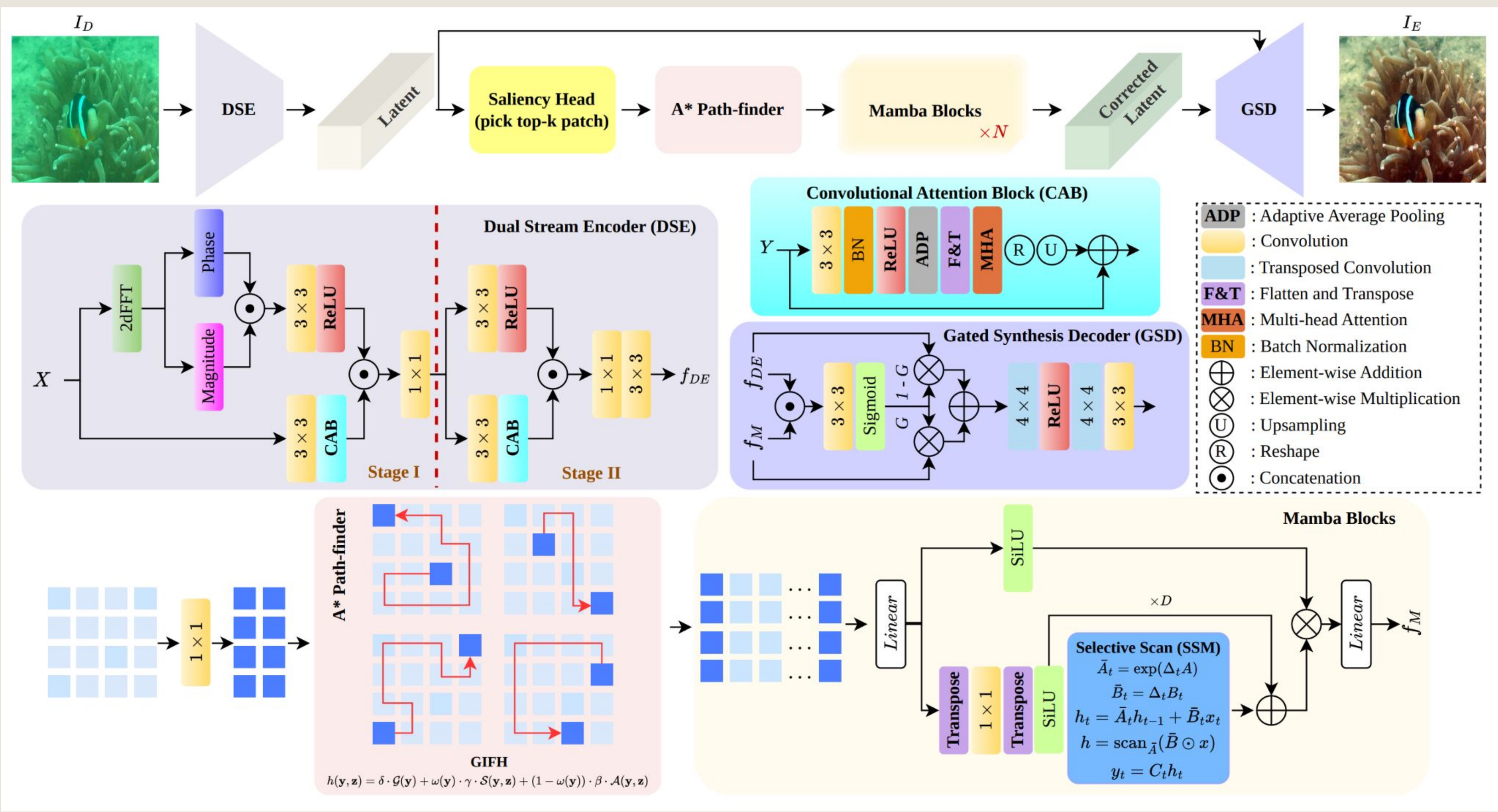

Figure 2: Overall Architecture of D2Mamba. The network extracts spatial and frequency features via a dual-stream encoder, adaptively traverses them using A* guided GIFH, enhances sequential features with Mamba blocks, and reconstructs images through a gated decoder.

Experimentations

	Methods	DWNet TOMM'23	U-Shape TIP'23	DM-Water MM'23	Unformer TAI'24	Spectroformer WACV'24	CE-VAE WACV'25	Phaseformer WACV'25	SS-UIE AAAI'25	Ours
	$\#Params(M)\downarrow$	0.48	65.6	10.00	7.31	2.40	80.74	1.77	4.25	**0.78**
	$Flops(G)\downarrow$	18.2	66.2	-	13.98	15.75	245.92	13.00	19.37	**7.06**
UIEB	SSIM ↑	0.843	0.910	0.909	0.861	0.917	0.870	0.928	0.909	**0.937**
	PSNR ↑	23.17	22.91	23.03	23.63	24.96	23.55	25.98	25.04	**25.70**
	UIQM ↑	4.237	4.330	2.908	4.238	4.200	4.323	4.418	4.441	**4.495**
	UISM ↑	7.236	7.296	4.058	7.271	7.312	7.299	7.132	7.086	**7.378**
LSUI	SSIM ↑	0.924	0.932	0.919	0.910	0.885	0.902	0.922	0.891	**0.935**
	PSNR ↑	28.98	26.13	29.56	28.88	25.08	29.65	29.60	28.61	**29.69**
	UIQM ↑	4.554	4.381	-	4.383	4.265	4.442	4.400	4.645	**4.669**
	UISM ↑	7.122	6.927	-	7.045	6.938	6.964	7.020	6.941	**7.144**
U45	UIQM ↑	4.549	3.997	3.854	4.133	3.921	4.007	4.245	4.250	**4.617**
	UISM ↑	6.873	6.962	7.120	7.005	7.088	6.959	7.099	7.073	**7.270**
	BRISQUE ↓	20.986	21.565	19.334	27.068	18.728	24.520	17.504	19.775	19.969
UCCS	UIQM ↑	4.476	4.201	4.301	4.436	4.229	4.352	3.980	4.278	**4.520**
	UISM ↑	6.988	6.913	7.008	6.843	6.898	6.887	6.974	6.869	**6.989**
	BRISQUE ↓	26.977	28.073	26.429	28.682	26.392	27.933	27.710	30.129	**25.832**
C60	UIQM ↑	4.227	4.514	4.184	4.204	4.158	4.375	4.286	4.417	**4.688**
	UISM ↑	7.438	6.764	7.421	7.299	7.488	7.353	7.499	7.422	**7.476**
	BRISQUE ↓	25.764	31.183	23.709	29.738	24.951	27.515	23.168	26.509	25.170
SQUID	UIQM ↑	2.434	2.266	-	-	2.558	1.965	2.335	2.135	**2.557**
	UISM ↑	7.373	7.275	-	-	7.193	7.189	7.357	7.302	**7.323**
	BRISQUE ↓	17.468	27.919	-	-	26.593	33.291	29.196	21.884	**11.489**

Table 1: Quantitative comparison of D2Mamba with existing methods on two full-reference datasets (UIEB, LSUI) and four no-reference datasets (U45, UCCS, C60, SQUID). The top-performing method is highlighted in red, the second-best in blue, and the third-best in green. (↑ higher is better, ↓ lower is better).

Scan Type	Latent Pixels	SSIM	PSNR
Raster	4096	0.926	25.30
Bidirectional	4096	0.928	25.37
Cross-scan	4096	0.929	25.43
Diagonal	4096	0.928	25.38
GIFH guided A^*	**110 - 125**	**0.937**	**25.70**

Table 2: Comparison of quantitative results on D2Mamba using various scanning types on the UIEB dataset.

L_2	L_{SSIM}	L_P	L_{SWAL}	SSIM	PSNR
✓	×	×	×	0.918	24.88
✓	✓	×	×	0.931	25.14
✓	✓	✓	×	0.934	25.26
✓	✓	✓	✓	**0.937**	**25.70**

Table 3: Comparison of quantitative results on D2Mamba obtained using different loss settings on the UIEB dataset.

$\#Paths$	$\#MamabaLayers$	SSIM	PSNR
8	2	0.937	25.68
4	**4**	**0.937**	**25.70**
2	8	0.932	25.27

Table 4: Comparison of quantitative results on D2Mamba by varying path numbers and Mamba layers on the UIEB dataset.

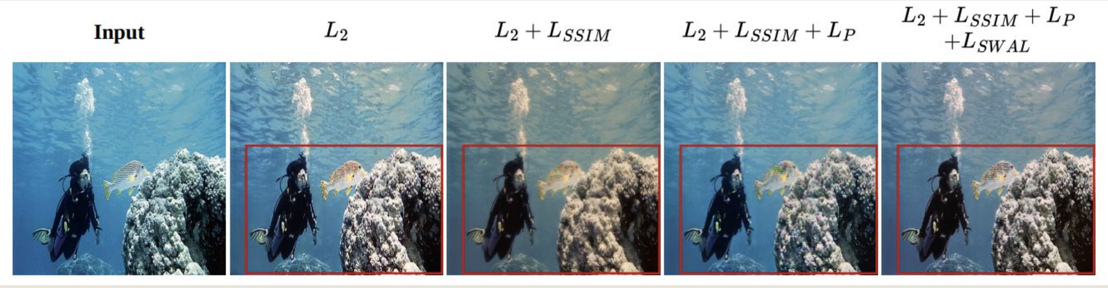

Figure 1: Enhanced image under various loss configuration.

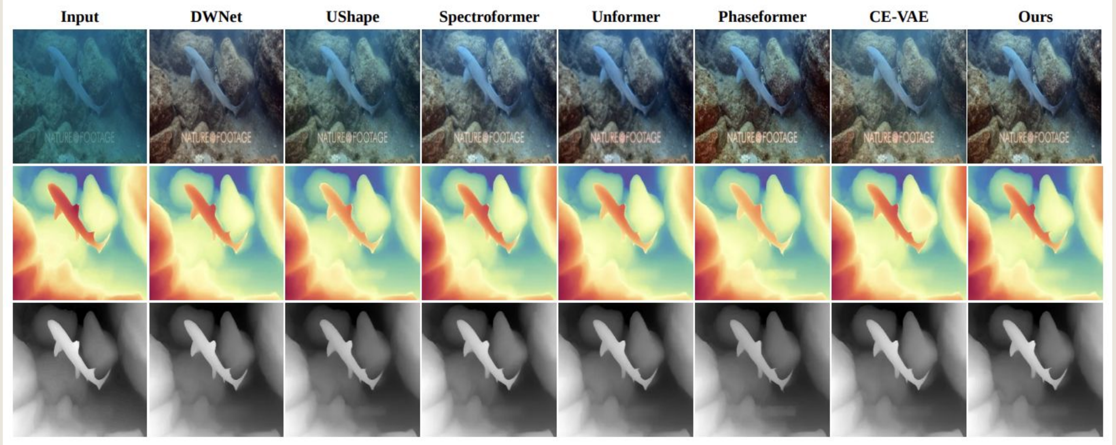

Figure 2: Depth map of degraded and enhanced images.

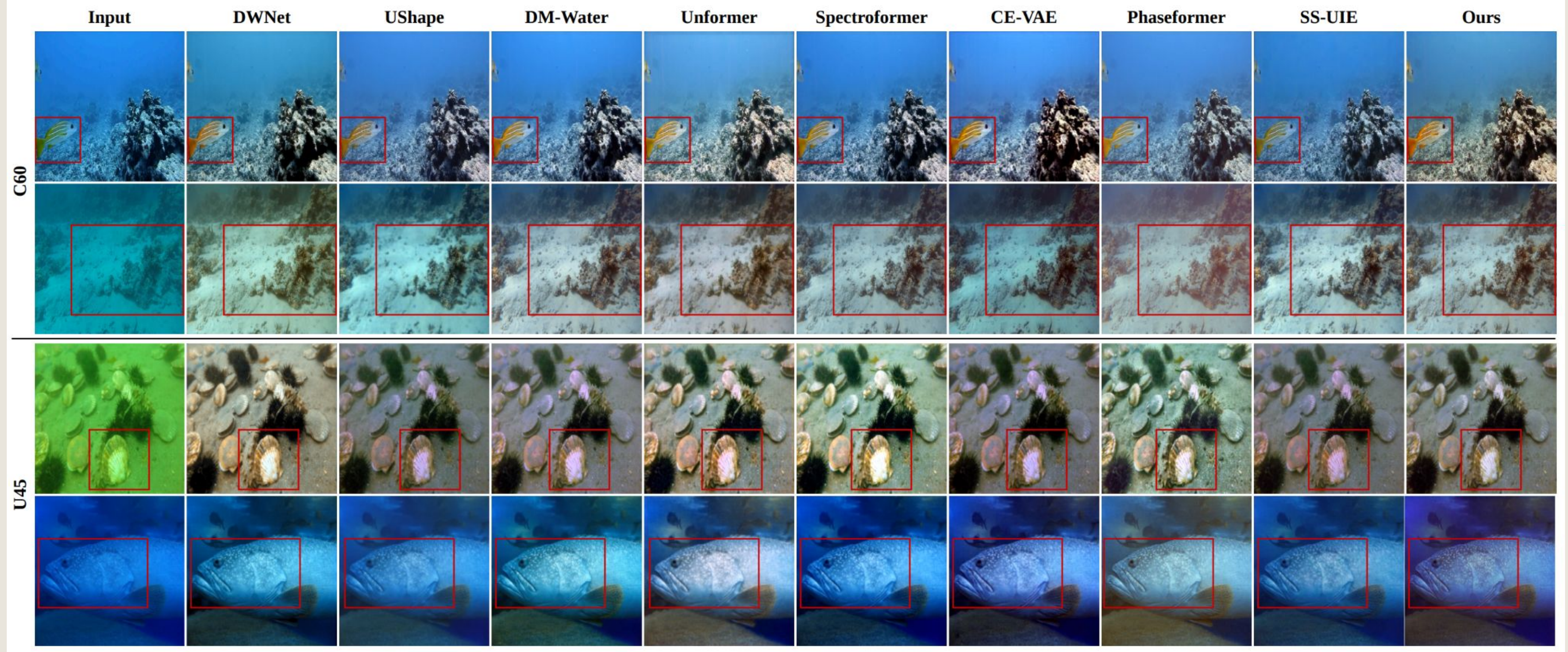

Figure 3: Qualitative results on C60 and U45. D2Mamba method produces clearer structures and more natural colors compared to other existing methods.

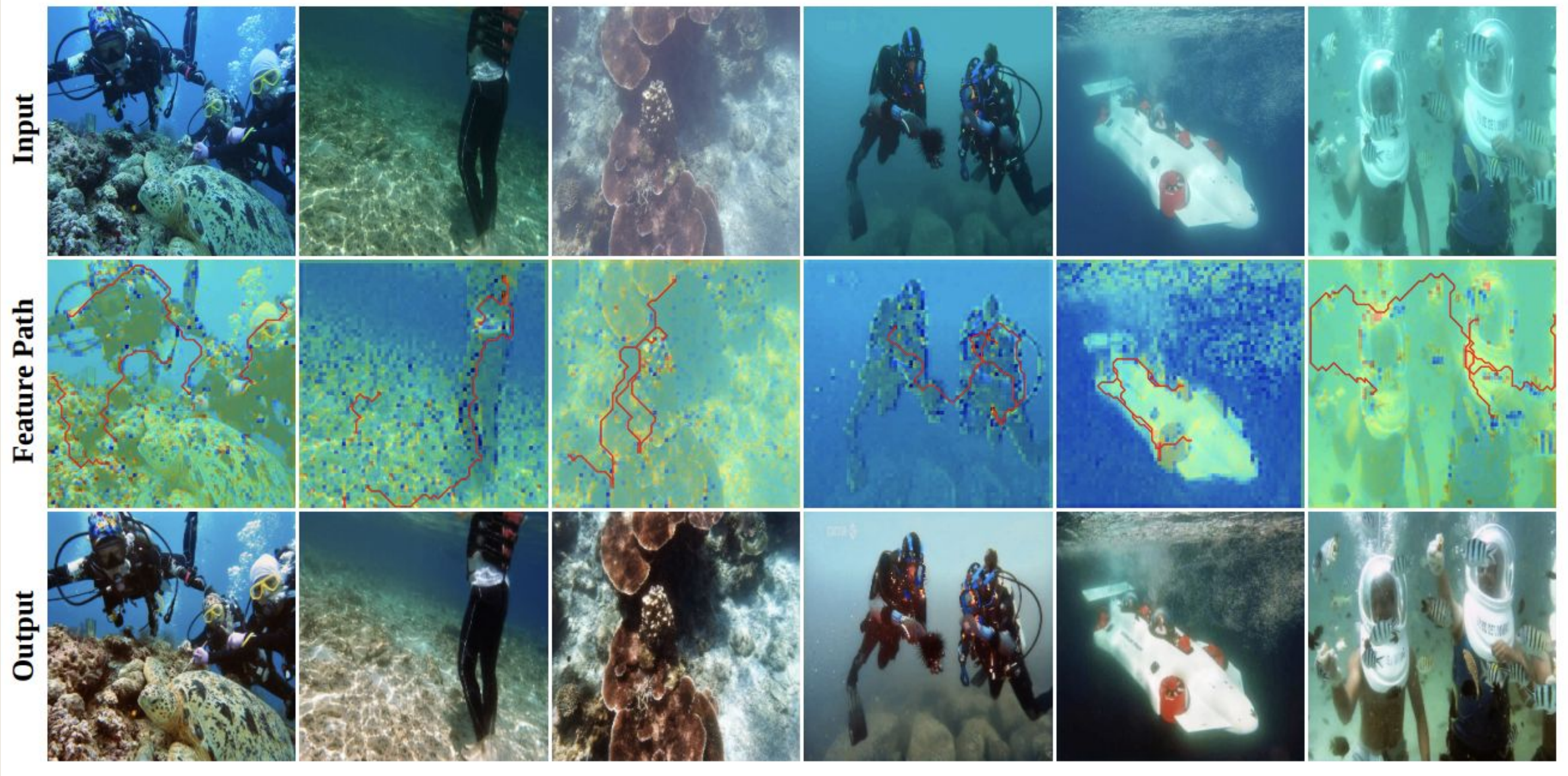

Figure 4: Visualization of learned paths using GIFH.

Conclusion

- Proposed D2Mamba, a lightweight Mamba-based framework for underwater image enhancement.
- Integrates dual-stream spatial-frequency encoding, geodesic-guided feature traversal, Mamba-based correction, and gated synthesis reconstruction to capture local details and global dependencies effectively.
- Outperforms existing UIE methods with lower computational requirements, achieving a strong balance of efficiency, robustness, and visual fidelity.

References

1. Prasen Sharma, Ira Bisht, and Arijit Sur. Wavelength-based attributed deep neural network for underwater image restoration. ACM Transactions on Multimedia Computing, Communications and Applications, 19(1):1–23, 2023
2. Lintao Peng, Chunli Zhu, and Liheng Bian. U-shape transformer for underwater image enhancement. IEEE transactions on image processing, 32:3066–3079, 2023.
3. Yuhao Qing, Yueying Wang, Huaicheng Yan, Xiangpeng Xie, and Zhengguang Wu. Unformer: A transformer-based approach for adaptive multi-scale feature aggregation in underwater image enhancement. IEEE Transactions on Artificial Intelligence, 2024.
4. Raqib Khan, Priyanka Mishra, Nancy Mehta, Shruti S Phutke, Santosh Kumar Vipparthi, Sukumar Nandi, and Subrahmanyam Murala. Spectroformer: Multi-domain query cascaded transformer network for underwater image enhancement. In Proceedings of the IEEE/CVF winter conference on applications of computer vision, pages 1454–1463, 2024.
5. Rita Pucci and Niki Martinel. Ce-vae: Capsule enhanced variational autoencoder for underwater image enhancement. In 2025 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV), pages 2113–2123. IEEE, 2025.
6. MD Raqib Khan, Anshul Negi, Ashutosh Kulkarni, Shruti S Phutke, Santosh Kumar Vipparthi, and Subrahmanyam Murala. Phaseformer: Phase-based attention mechanism for underwater image restoration and beyond. In 2025 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV), pages 9618–9629. IEEE, 2025.
7. Lintao Peng and Liheng Bian. Adaptive dual-domain learning for underwater image enhancement. In Proceedings of the AAAI Conference on Artificial Intelligence, pages 6461– 6469, 2025.

Contact

Alik Pramanick
Soumajit Roy
Arijit Sur