AAAI-26 / IAAI-26 / EAAI-26
JANUARY 20–27, 2026 | SINGAPORE

Hero-Mamba: Mamba-based Dual Domain Learning for Underwater Image Enhancement

Tejeswar Pokuri, Shivarth Rai
Manipal Academy of Higher Education, India

MANIPAL
ACADEMY *of* HIGHER EDUCATION
(Deemed to be University under Section 3 of the UGC Act, 1956)
INSPIRED BY LIFE

Introduction

$$I_c = J_c . t_c + B_c(1 - t_c)$$

- I_c is the observed degraded image
- J_c is the clear image
- t_c is the transmission map
- B_c is the background light

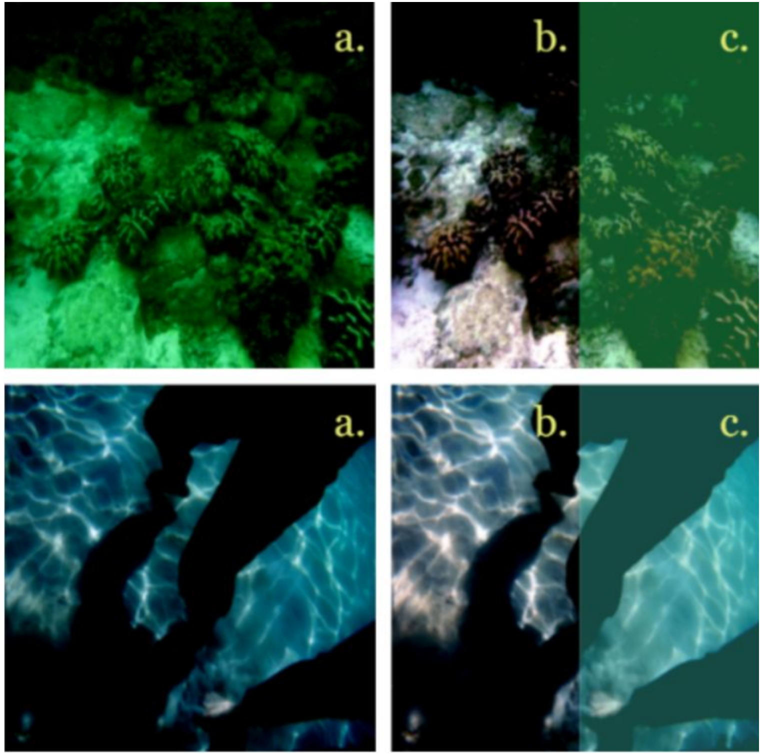


Fig 1. (Right) Visualizing the contribution of the background light prior on UIEB samples. (a) Original degraded images. (b) Ground Truth. (c) The background light prior, as estimated by our ColorFusion block, overlaid for visualization. The stark color difference between (b) and (c) highlights the severe color degradation caused by background light.

[1] Jaffe, J. S. 2002. Computer modeling and the design of optimal underwater imaging systems. IEEE journal of oceanic engineering, 15(2): 101–111.

Related Works

Physical Models: Methods like UDCP[1] and Haze-Lines often lack generalization across diverse water types.

CNN-based: Methods like Ucolor[2] use multi-color space embeddings but lack long-range dependency modeling.

Transformer-based: Methods like U-Shape Transformer[3] offer global context but are computationally expensive.

Why State Space Models (SSMs)?

Mamba Architecture: Introduces selective scan mechanisms to model long-range dependencies with **linear complexity O(N)**.

Advantage: Overcomes the computational bottleneck of Transformers while maintaining the global receptive field capability needed for UIE.

Recent Adoption: Emerging methods like SS-UIE[4] and WaterMamba[5] utilize SSMs for efficient spatial-spectral learning.

[1] Drews Jr, P.; do Nascimento, E.; Moraes, F.; Botelho, S.; and Campos, M. 2013. Transmission Estimation in Underwater Single Images. In 2013 IEEE International Conference on Computer Vision Workshops, 825–830.
[2] Li, C.; Anwar, S.; Hou, J.; Cong, R.; Guo, C.; and Ren, W. 2021. Underwater Image Enhancement via Medium Transmission-Guided Multi-Color Space Embedding. IEEE Transactions on Image Processing, 30: 4985–5000.
[3] L. Peng, C. Zhu and L. Bian, "U-Shape Transformer for Underwater Image Enhancement," in IEEE Transactions on Image Processing, vol. 32, pp. 3066-3079, 2023, doi: 10.1109/TIP.2023.3276332.
[4] Peng, L.; and Bian, L. 2025. Adaptive Dual-domain Learning for Underwater Image Enhancement. Proceedings of the AAAI Conference on Artificial Intelligence, 39(6): 6461–6469.
[5] Guan, M.; Xu, H.; Jiang, G.; Yu, M.; Chen, Y.; Luo, T.; and Song, Y. 2024. WaterMamba: Visual State Space Model for Underwater Image Enhancement. arXiv:2405.08419.

Our Contributions

Hero Mamba Architecture

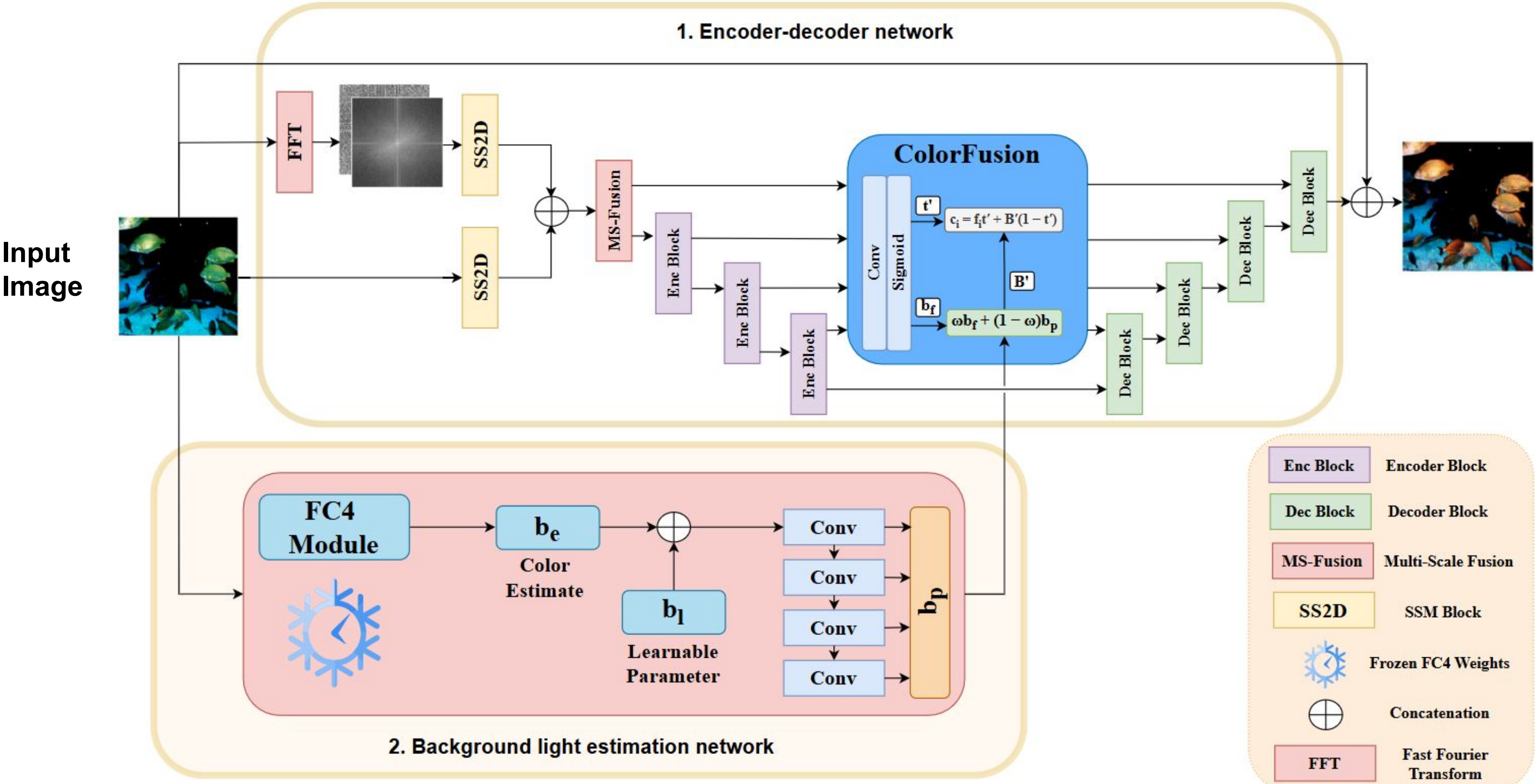
- Mamba-based dual-domain learning framework
- jointly processes spatial (RGB) and spectral (FFT) features to efficiently model global long-range dependencies
- Physics Guided Background Color Restoration.

Hero-Mamba is compared against **10 state-of-the-art methods** on the **LSUI** and **UIEB** benchmarks using SSIM, PSNR, FSIM, and LPIPS metrics.

Results

- **SSIM of 0.913 and PSNR of 25.802 on LSUI dataset,** outperforming all the existing state of the art models.
- **SSIM of 0.942 on UIEB dataset**, at par with existing state of the art models.
- Cross-dataset study for validating generalizability: **SSIM - 0.868** (LSUI → UIEB) and **SSIM - 0.846** (UIEB → LSUI).

Overview of Hero-Mamba Architecture



1. Encoder-decoder network

Input Image

FFT

SS2D

SS2D

MS-Fusion

Enc Block

Enc Block

Enc Block

ColorFusion

Conv

Sigmoid

t'

$c_i = f_i t' + B'(1 - t')$

B'

b_f

$\omega b_f + (1 - \omega) b_p$

Dec Block

Dec Block

Dec Block

Dec Block

2. Background light estimation network

FC4 Module

b_e

Color Estimate

b_l

Learnable Parameter

Conv

Conv

Conv

Conv

b_p

Enc Block	Encoder Block
Dec Block	Decoder Block
MS-Fusion	Multi-Scale Fusion
SS2D	SSM Block
❄	Frozen FC4 Weights
⊕	Concatenation
FFT	Fast Fourier Transform

Fig 2. Architectural design of Hero-Mamba, utilizing spatial and spectral domains for accurate feature reconstruction, and ColorFusion block for enhanced color restoration. Using SS2D layers allows for long-range feature learning efficiently.

ColorFusion Block

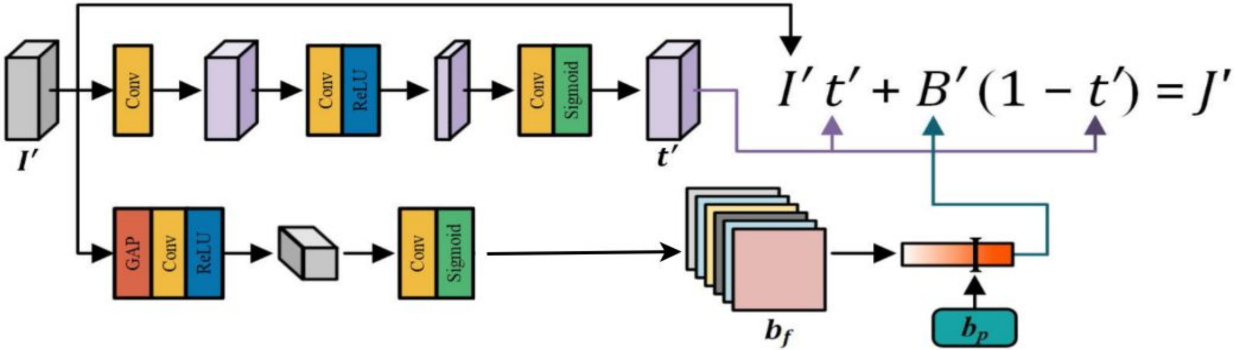


Fig 3. Architectural Diagram for Color Fusion Block.

Step 1 : B’ Computation

$$B' = \omega . b_f + (1 - \omega) . b_p$$

Step 2 : Transmission Map Computation

$$t' = \mathit{Sigmoid}(\mathit{Conv}(I'))$$

Step 3: Color Fusion Feature Computation

$$c_i = I' . t' + B' (1 - t')$$

MS-Fusion Block

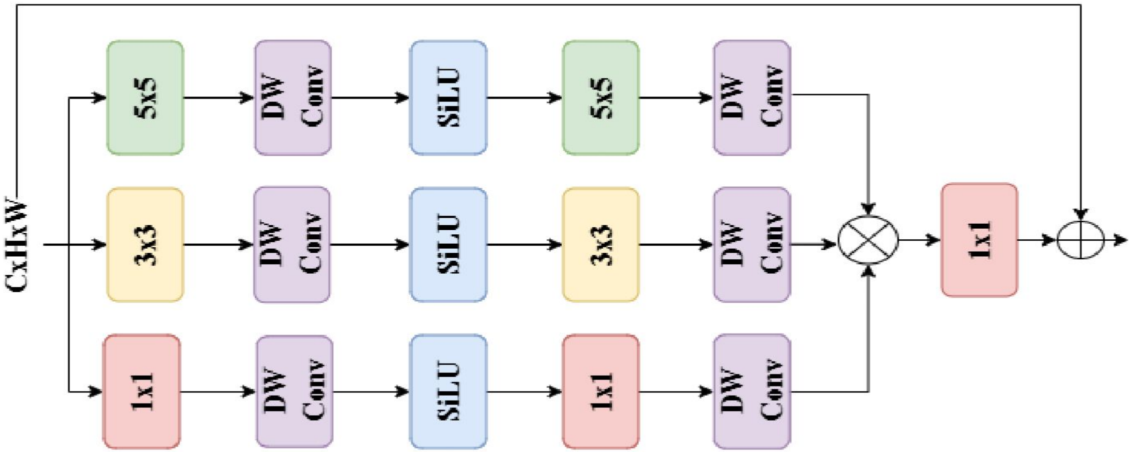


Fig 4. Architectural Diagram for MS-Fusion Block.

Architecture: Utilizes parallel convolution branches with varying kernel sizes ***(1x1, 3x3, 5x5)*** [1].

Goal: Captures features at multiple scales simultaneously to restore both fine local details and broader structural edges.

Multi-Scale Fusion: Features from all branches are concatenated and fused back with the input to provide a comprehensive representation.

[1] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, “Going Deeper with Convolutions,” *arXiv preprint arXiv:1409.4842*, 2014.

Qualitative and Quantitative Analysis - UIEB

Table 1. Quantitative comparison on UIEB Dataset. The best result is highlighted in red, and the second-best result is in blue.

Paper Name	**Venue**	**SSIM** ↑	**PSNR** ↑	**LPIPS** ↓	**FSIM** ↑
U-Gan	ICRA 2018	0.805	19.676	0.197	0.912
FUnIE-Gan	RAL 2020	0.814	18.781	0.163	0.919
Semi-UIR	CVPR 2023	0.821	23.400	0.157	0.932
PUIE-Net	ECCV 2022	0.854	21.501	0.132	0.863
WF-Diff	CVPR 2024	0.873	27.260	0.139	0.897
NU2Net	AAAI 2023	0.907	22.633	0.100	0.949
Water Mamba	Arxiv 2024	0.931	24.751	0.143	0.973
Hero-Mamba (Ours)		0.942	24.526	0.125	0.945

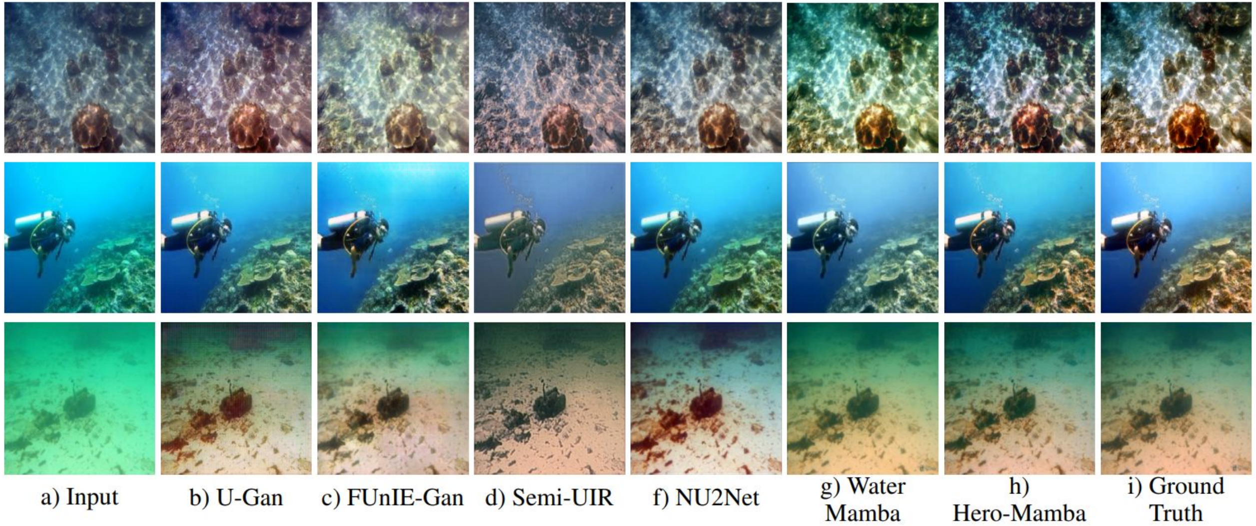
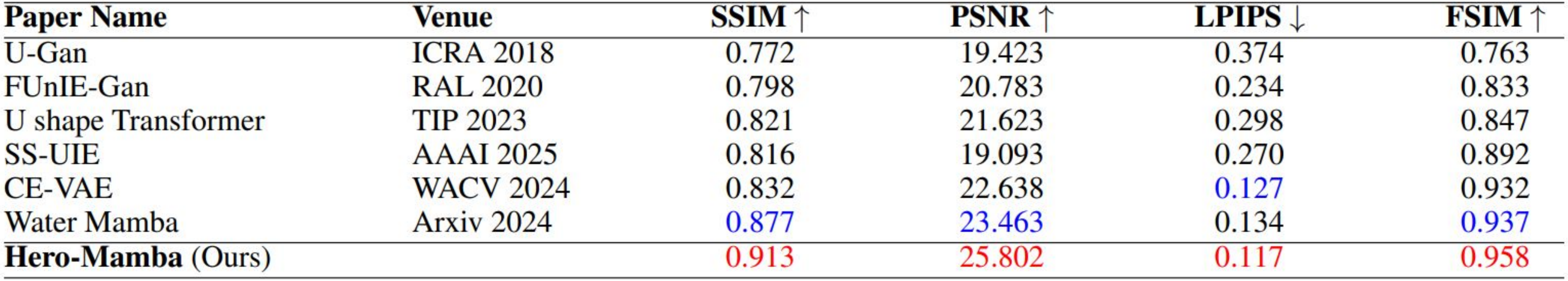


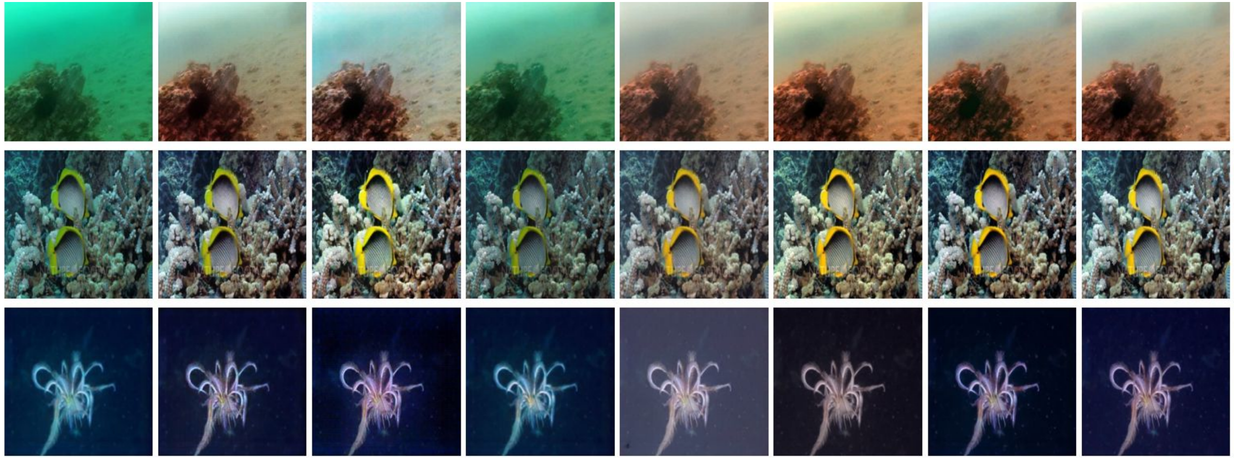
Fig 5. Visual comparison of enhancement results by various models on UIEB dataset.

Qualitative and Quantitative Analysis - LSUI

Table 2. Quantitative comparison on LSUI Dataset. The best result is highlighted in red, and the second-best result is in blue.



Paper Name	**Venue**	**SSIM ↑**	**PSNR ↑**	**LPIPS ↓**	**FSIM ↑**
U-Gan	ICRA 2018	0.772	19.423	0.374	0.763
FUnIE-Gan	RAL 2020	0.798	20.783	0.234	0.833
U shape Transformer	TIP 2023	0.821	21.623	0.298	0.847
SS-UIE	AAAI 2025	0.816	19.093	0.270	0.892
CE-VAE	WACV 2024	0.832	22.638	0.127	0.932
Water Mamba	Arxiv 2024	0.877	23.463	0.134	0.937
Hero-Mamba (Ours)		0.913	25.802	0.117	0.958

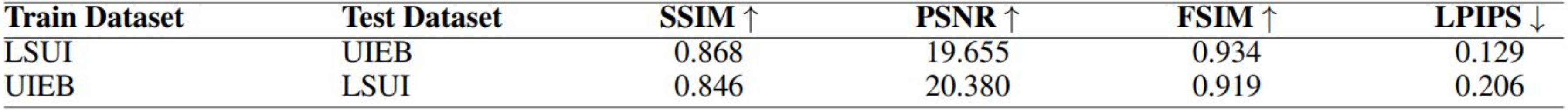


a) Input b) U-Gan c) FUnIE-Gan d) SS-UIE e) CE-VAE f) Water Mamba g) Hero-Mamba h) Ground Truth

Fig 6. Visual comparison of enhancement results by various models on LSUI dataset.

Cross Dataset Study

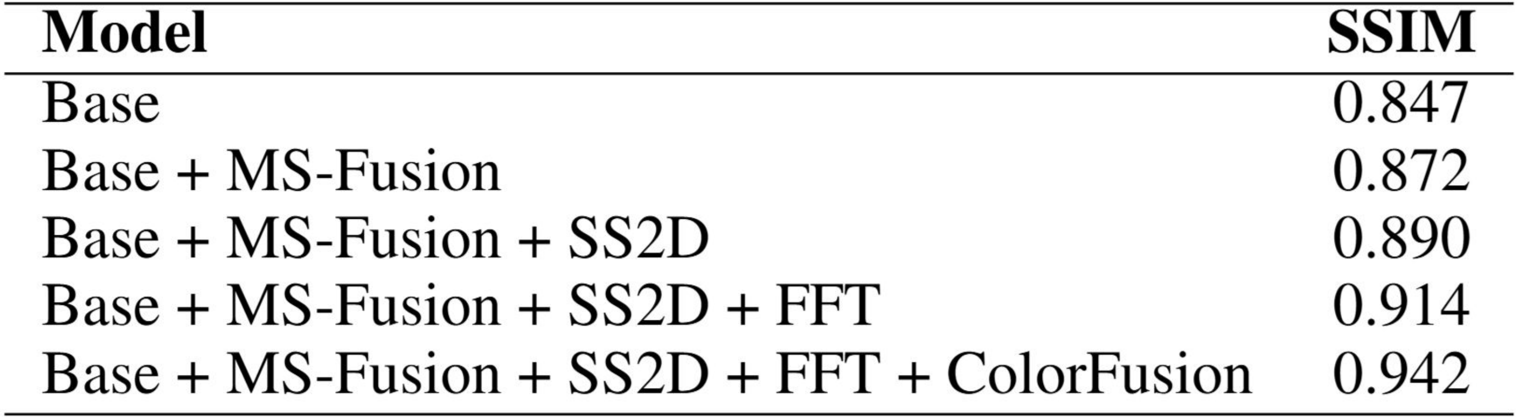
Table 3. Cross-dataset study results for Hero-Mamba.



Train Dataset	Test Dataset	SSIM ↑	PSNR ↑	FSIM ↑	LPIPS ↓
LSUI	UIEB	0.868	19.655	0.934	0.129
UIEB	LSUI	0.846	20.380	0.919	0.206

Ablation Study

Table 4. Break-down ablation study for Hero-Mamba



Model	SSIM
Base	0.847
Base + MS-Fusion	0.872
Base + MS-Fusion + SS2D	0.890
Base + MS-Fusion + SS2D + FFT	0.914
Base + MS-Fusion + SS2D + FFT + ColorFusion	0.942

Thank you for Listening!

Contact us at:
Shivarth Rai
{raishivarth@gmail.com}
Tejeswar Pokuri
{tejeswarpokuri3@gmail.com}