Frequency Modulated Deformable Transformer for Underwater Image Enhancement

Supplementary Material

Overview of Supplementary Material

The supplementary material includes:

1. Significance of Spatio-channel Attentive Offset Extractor

There has been extensive research on the ability of deformable convolution. However, an important concern is the presence of irrelevant features as a result of offsets that exceed the contextually relevant regions of the image. To address this, we introduce a spatio-channel attentive offset extractor (*please refer Figure S 1*) which is sensitive to fine structural variations and color shifts induced by underwater conditions, and refines the offset extraction process by aligning it with color variations, ultimately enhancing both structural details and color accuracy in the restored images. The proposed spatio-channel attentive offset employs both spatial and channel attention on features to calculate offsets. This approach ensures that deformable convolution concentrate on the most relevant regions of the image and the color correlation, facilitated by structural and color-relevant offset cues. According to the feature map shown in Figure S 2, Qualitative results in Figure S 3 and the non-reference color metrics in Table S 4, the proposed spatio-channel attentive offset is effective in enhancing fine structural variations as well as generating realistic colors.

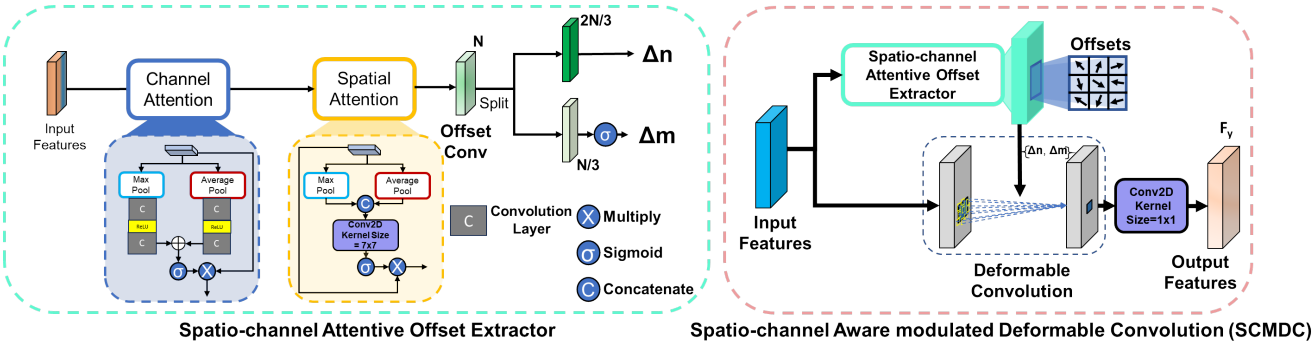

Figure S 1. The proposed spatio-channel attentive offset extractor and spatio-channel aware deformable convolution are illustrated in detail.

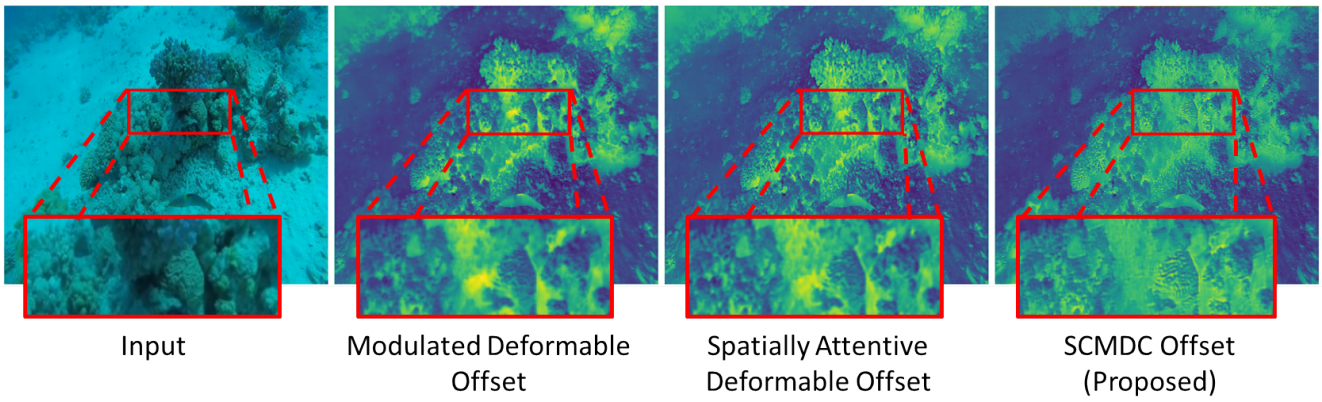

Figure S 2. Feature map visualization of various combinations of offset extractor. The proposed SCMDC offset extractor can extract more local spatial information (*as shown in the red box*) than the modulated deformable offset and spatially attentive deformable offset extractor, resulting in a superior structural variation in the proposed method output images.

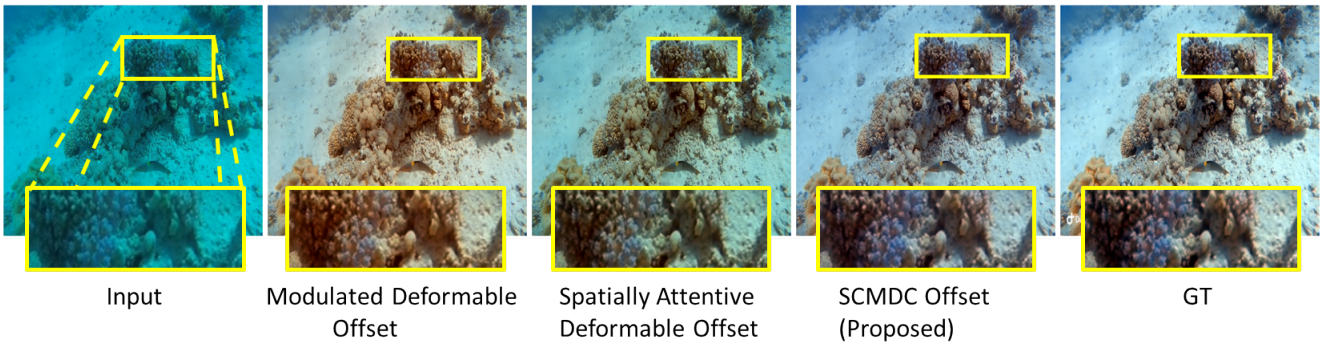

Figure S 3. Qualitative comparison for various offset extractors

2. Application of Underwater Image Enhancement

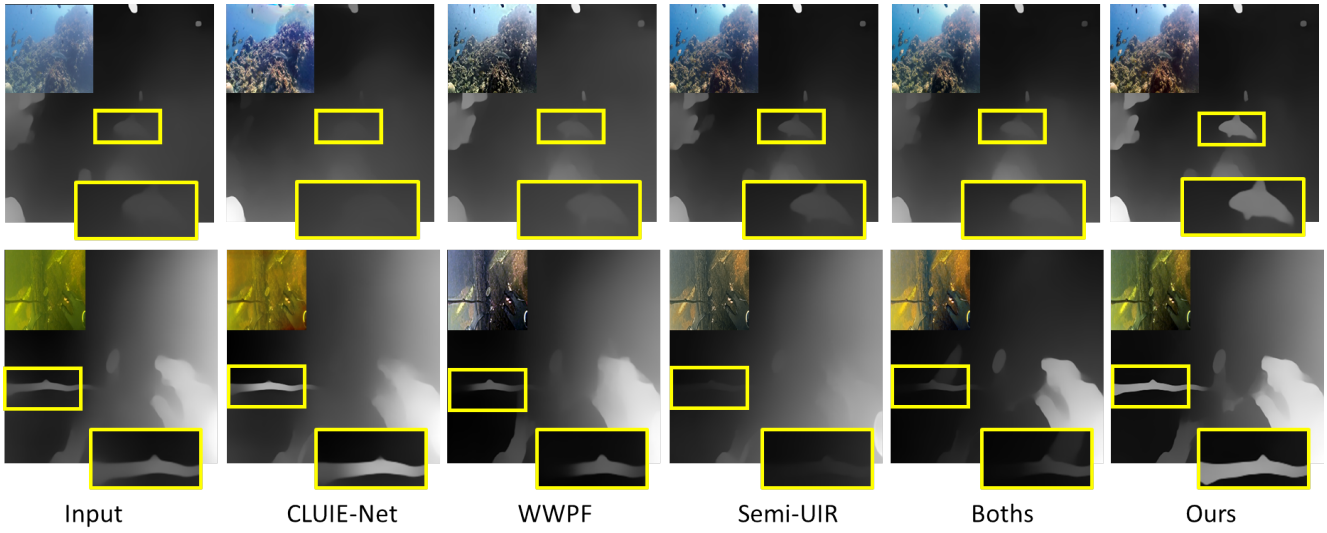

Figure S 4. Depth estimation analysis for degraded and enhanced images by the existing CLUIE [9], WWPF [16], Semi-UIR [4] and Boths [11], and the proposed method (ours).

The ultimate goal of every underwater enhancement method is its applicability to complex underwater scenarios. Hence, we have analysed our proposed method on high-level computer vision application like depth estimation. We have first passed the original degraded image through depth estimator and then we enhanced the image using various comparative methods CLUIE [9], WWPF [16], Semi-UIR [4], Boths [11] and proposed method and these enhanced images are passed through depth estimator. The depth estimation results in Figure S 4 depict that underwater image enhancement using proposed method is more accurate for application than other comparative methods. Consequently, proposed method validates its applicability to higher level computer vision task in complex underwater scenarios.

3. Analysis of Loss Functions

In the proposed work, we have utilized several loss functions to optimize the network. The network is trained with the content loss function $(\mathbb{L}_1)$. Further to reduce the difference in frequency space, we have employed FFT loss $(\mathbb{L}_F)$ [3] which calculates the similarity between the ground truth and output of the network as:

$$\mathbb{L}_F = \sum_{k=1}^{k} \parallel F(G_t) - F(O) \parallel_1 \tag{1}$$

where, $F(.)$ is the FFT of an image, G_t and O represent ground-truth and output respectively. Furthermore to maintain the feature level textural and structural similarity, the perceptual loss (L_p) is calculated using the VGG-16 pre-trained module as:

$$\mathbb{L}_p = \sum_{s=1}^{s} \parallel \psi_s(G_t) - \psi_s(O) \parallel_1 \tag{2}$$

where, ψ_s is pre-trained model of VGG-16 $(s \in (1, S))$. Also, the contrastive loss $(\mathbb{L}_C)$ is calculated to maximize and minimize the difference between input-output and output-groudtruth respectively as:

$$\mathbb{L}_C(I, O, G_t) = \sum_{s=1}^{s} \frac{\parallel \psi_s(O) - \psi_s(G_t) \parallel_1}{\parallel \psi_s(I) - \psi_s(O) \parallel_1} \tag{3}$$

Therefore, the total loss is represented as:

$$\mathbb{L}_{total} = \lambda_1 \mathbb{L}_1 + \lambda_2 \mathbb{L}_F + \lambda_3 \mathbb{L}_p + \lambda_4 \mathbb{L}_C \tag{4}$$

we set weights as $\lambda_1 = 1, \lambda_2 = 0.1, \lambda_3 = 15$ and $\lambda_4 = 5$. The various combinations of loss function play significant importance in optimizing the network. Table S 1 depicts the quantitative analysis on importance of various combinations of loss functions on UIEB [7] dataset.

Table S 1. Quantitative analysis on various loss function combinations.

Loss Setting	PSNR	SSIM
$\mathbb{L}_1$	23.21	0.921
$\mathbb{L}_1 + \mathbb{L}_F$	24.20	0.933
$\mathbb{L}_1 + \mathbb{L}_F + \mathbb{L}_p$	24.75	0.933
$\mathbb{L}_1 + \mathbb{L}_F + \mathbb{L}_p + \mathbb{L}_C$	**25.79**	**0.955**

4. Analysis on More Real-world Dataset

UFO-120 dataset [5]: We used 120 natural underwater images collected from oceanic explorations in multiple location presented in this dataset for qualitative and non-reference-based evaluation.

U-45 dataset [8]: The U45 is sorted into three subsets of the green, blue, and haze-like categories, where the subsets correspond to the color casts, low contrast and haze like effects of underwater degradation this dataset is consist of 45 real-world underwater images.

The Underwater Image Quality Measurement (UIQM) [13], Underwater Color Image Quality Evaluation (UCIQE) [15], Underwater Image Colourfulness Measure (UICM) [13], and Natural Image Quality Evaluator (NIQE) [12] metrics are used to evaluate the performance of proposed method over the LANET [10], Wave Net [14], CLUIE [9], WWPF [16], Semi-UIE [4], and Boths [11] methods on UFO-120 and U-45 dataset.

- The quantitative results on UFO-120 dataset are provided in Table S 2.
- The quantitative results on U-45 dataset are provided in Table S 3.

Quantitative results on UFO-120 and U-45 real-world dataset provided in Table S 2 and S 3 shows the reliability of the proposed network to perform significantly better on various dataset

Table S 2. Quantitative comparison of the proposed method and state-of-the-art methods on the UFO-120 [5] dataset for underwater image enhancement (Note: ↓: *lower is better*, ↑: *higher is better*).

Method	Publication	UIQM↑	UCIQE↑	UICM↑	NIQE↓
LANet [10]	RAL-22	3.5746	31.7775	-24.5989	4.9333
WaveNet [14]	TIP- 23	3.2742	33.0744	-26.6580	4.2037
CLUIE [9]	JOE-23	3.5780	33.1736	-21.7314	**2.2204**
WWPF [16]	TIP-23	3.5128	32.4230	-18.8301	4.3631
Semi-UIE [4]	CVPR-23	3.8571	31.7171	-18.9118	4.9278
Boths [11]	GRSL-23	3.6455	32.2657	-25.3624	4.8735
Ours		**3.965**	**33.2213**	**-11.0925**	4..8334

Table S 3. Quantitative comparison of the proposed method and state-of-the-art methods on the U45 [8] dataset for underwater image enhancement (Note: ↓: *lower is better*, ↑: *higher is better*)..

Method	Publication	UIQM↑	UCIQE↑	UICM↑	NIQE↓
LANet [10]	RAL-22	3.7598	29.9278	-18.3784	4.4656
WaveNet [14]	TIP- 23	4.2175	31.5389	-23.7600	3.8833
CLUIE [9]	JOE-23	4.1502	30.1410	-29.9435	3.7912
WWPF [16]	TIP-23	4.4037	31.0950	**-11.9490**	**3.7890**
Semi-UIE [4]	CVPR-23	4.5038	31.6943	-18.7717	4.4932
Boths [11]	GRSL-23	3.8110	31.5353	-24.6044	4.5535
Ours		**4.7426**	**31.8263**	-25.2993	4.4855

5. Quantitive Analysis in terms of Non-reference Metrics on UIEB and EUVP Dataset

We have used the existing LA-Net [10], FGan [6], Wave Net [14], CLUIE [9], WWPF [16], Semi-UIE [4], and Boths [11] methods for comparison in terms of UIQM [13], UCIQE [15], UICM [13], and NIQE [12] non reference metrics. For non-reference evaluation, only outputs of respective methods are considered.

- The Quantitative results on UIEB dataset is provided in Table S 4.
- The Quantitative results on EUVP dataset is provided in Table S 5.

The quantitative comparison in terms of non reference metrics shows the significant performance of the proposed method over the state-of-the-art approaches.

Table S 4. Quantitative comparison of the proposed method and state-of-the-art methods on the UIEB [7] dataset for underwater image enhancement (Note: ↓: *lower is better*, ↑: *higher is better*)..

Method	Publication	UIQM↑	UCIQE↑	UICM↑	NIQE↓
LA-Net [10]	RAL-22	3.9606	29.4789	-16.2530	4.1359
WaveNet [14]	TIP- 23	4.1778	31.8141	-20.1638	4.2404
CLUIE [9]	JOE-23	3.9181	30.9827	-18.8541	3.9824
WWPF [16]	TIP-23	4.1019	31.6688	-10.3432	**3.6050**
Semi-UIE [4]	CVPR-23	4.0392	31.6316	-15.8159	4.2791
Boths [11]	GRSL-23	3.7840	31.3833	-20.3492	4.1940
Ours		**4.4685**	**32.1202**	**-9.4388**	3.9400

Table S 5. Quantitative comparison of the proposed method and state-of-the-art methods on the EUVP [6] dataset for underwater image enhancement (Note: ↓: *lower is better*, ↑: *higher is better*).

Method	Publication	UIQM↑	UCIQE↑	UICM↑	NIQE↓
F-Gan [6]	RAL-20	4.0860	31.2317	-5.4831	5.08045
LANet [10]	RAL-22	3.6507	31.4194	-20.9812	5.0746
WaveNet [14]	TIP- 23	3.6569	28.9556	-16.3435	5.1232
Boths [11]	GRSL-23	3.9358	**31.4412**	-9.4974	5.1258
Ours		**4.0951**	30.0769	**-3.0080**	**5.0390**

6. Qualitative Results on Synthetic and Real-world Datasets

In this section, we provided more qualitative results on synthetic (UIEB [7], EUVP [6], UFO-120 [5]) and Real-world (U-45 [8], Color-checker [2]) datasets for underwater image enhancement.

- The qualitative results on sea-thru [1] dataset is provided in Figure S 5
- The qualitative results on UFO-120 [5] dataset is provided in Figure S 6
- The qualitative results on U-45 [8] dataset is provided in Figure S 7
- The qualitative results on UIEB [7] dataset is provided in Figure S 8
- The qualitative results on EUVP [6] dataset is provided in Figure S 9
- The qualitative results on Color-checker [2] dataset is provided in Figure S 10

The visual result analysis shows that the significant performance of the proposed method over the state-of-the-art approaches.

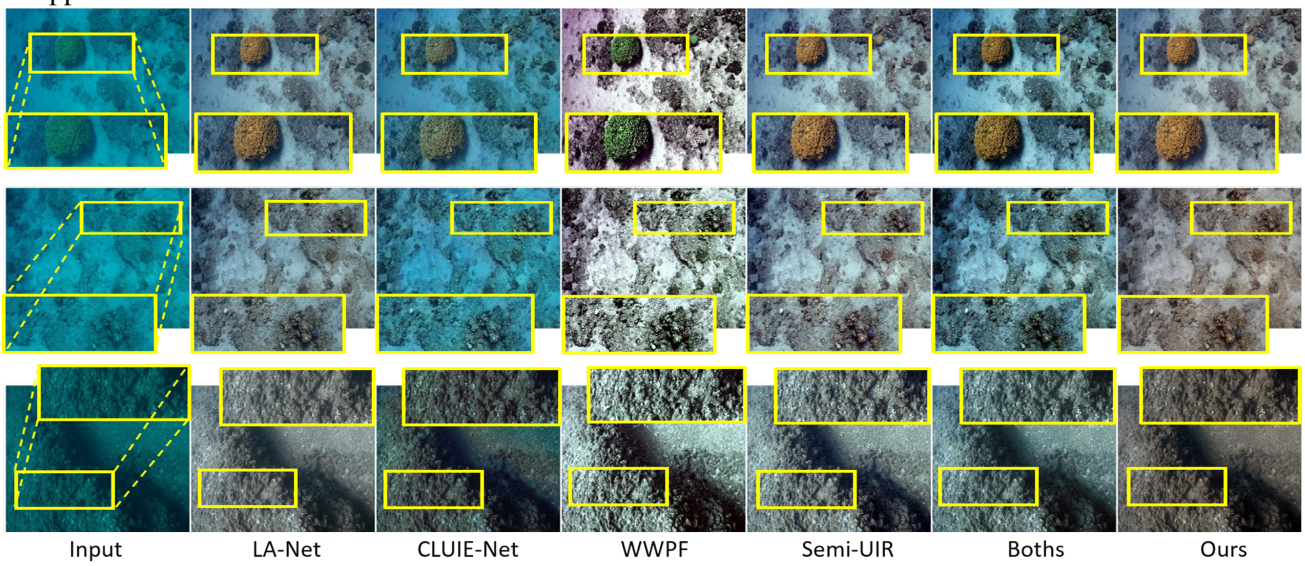

Figure S 5. Qualitative results comparison on Sea-thru dataset with state-of-the-art (LA-Net [10], WaveNet [14], CLUIE [9], WWPF [16], Semi-UIE [4] and Boths [11]) methods for underwater image enhancement.

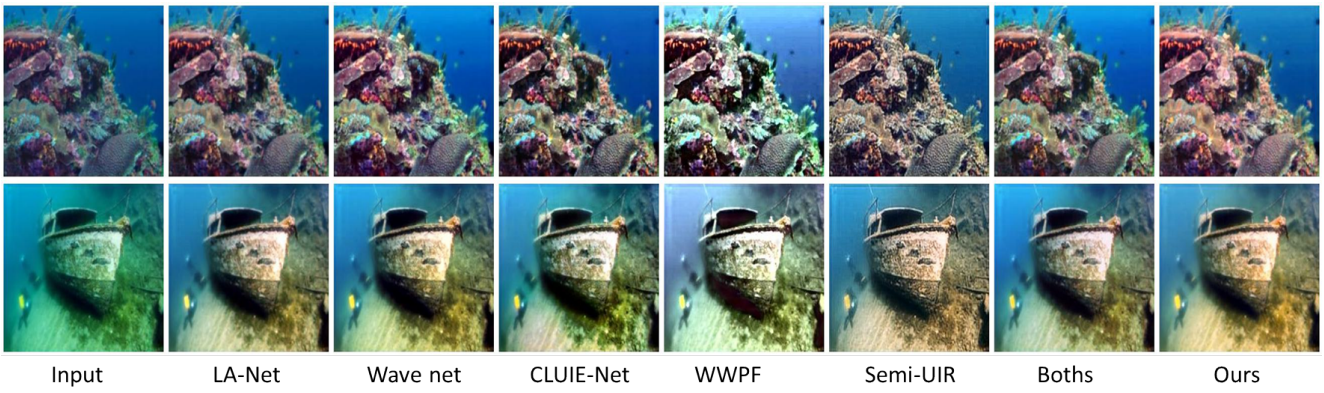

Figure S 6. Qualitative results comparison on UFO-120 dataset with state-of-the-art (LA-Net [10], WaveNet [14], CLUIE [9], WWPF [16], Semi-UIE [4] and Boths [11]) methods for underwater image enhancement.

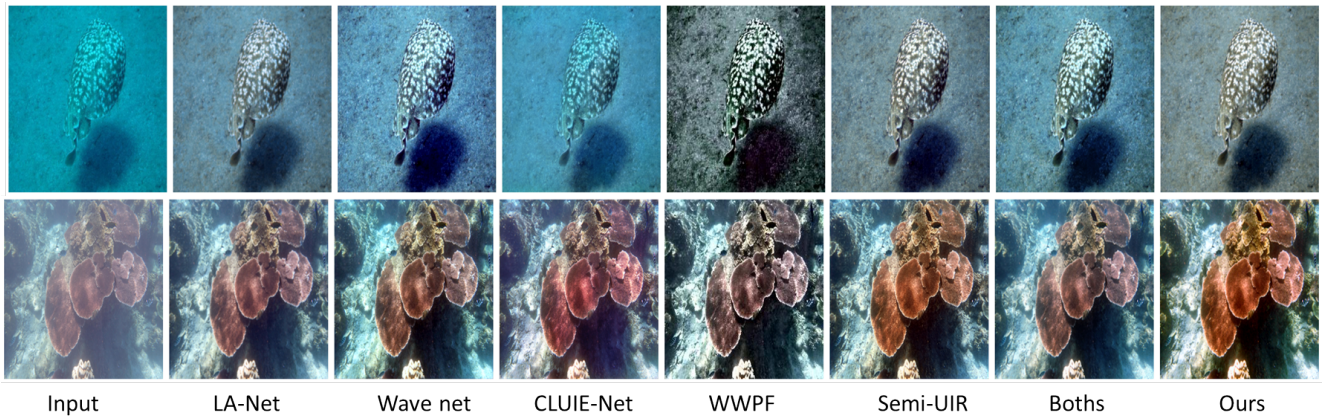

Figure S 7. Qualitative results comparison on U-45 dataset with state-of-the-art (LA-Net [10], WaveNet [14], CLUIE [9], WWPF [16], Semi-UIE [4] and Boths [11]) methods for underwater image enhancement.

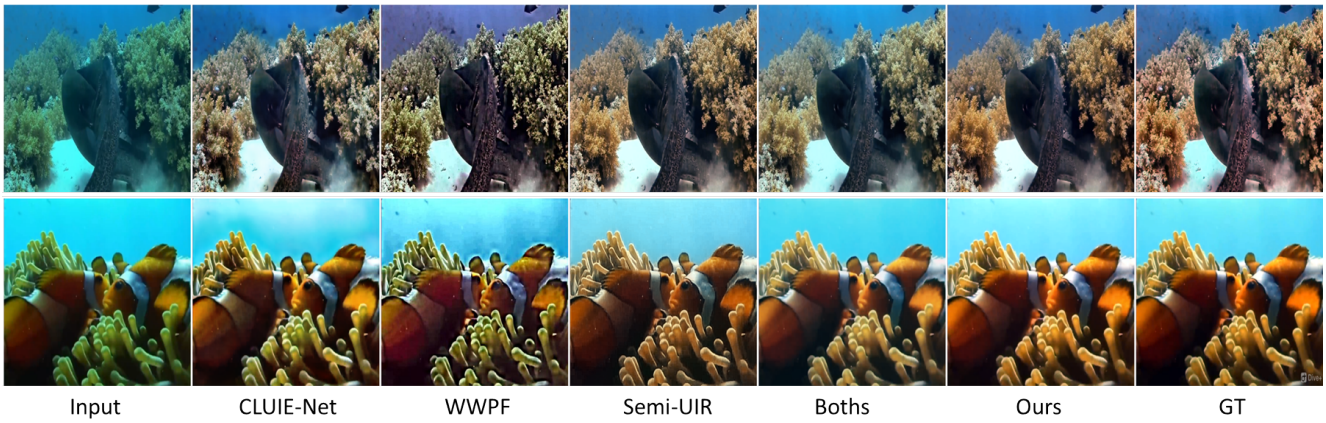

Figure S 8. Qualitative results comparison on UIEB dataset with state-of-the-art (CLUIE [9], WWPF [16], Semi-UIE [4] and Boths [11]) methods for underwater image enhancement.

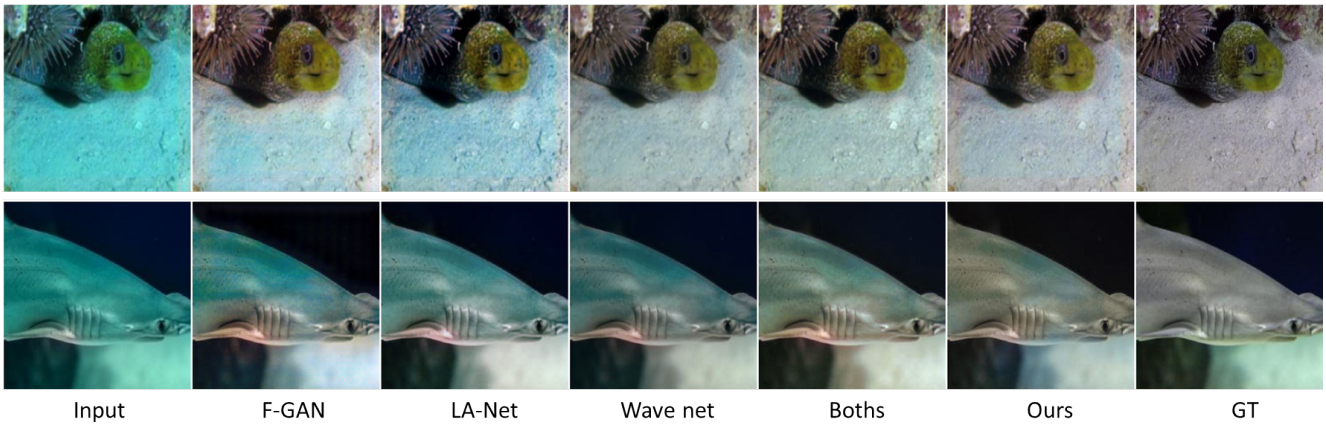

Figure S 9. Qualitative results comparison on EUVP dataset with state-of-the-art (F-Gan [6], LA-Net [10], WaveNet [14], and Boths [11]) methods for underwater image enhancement.

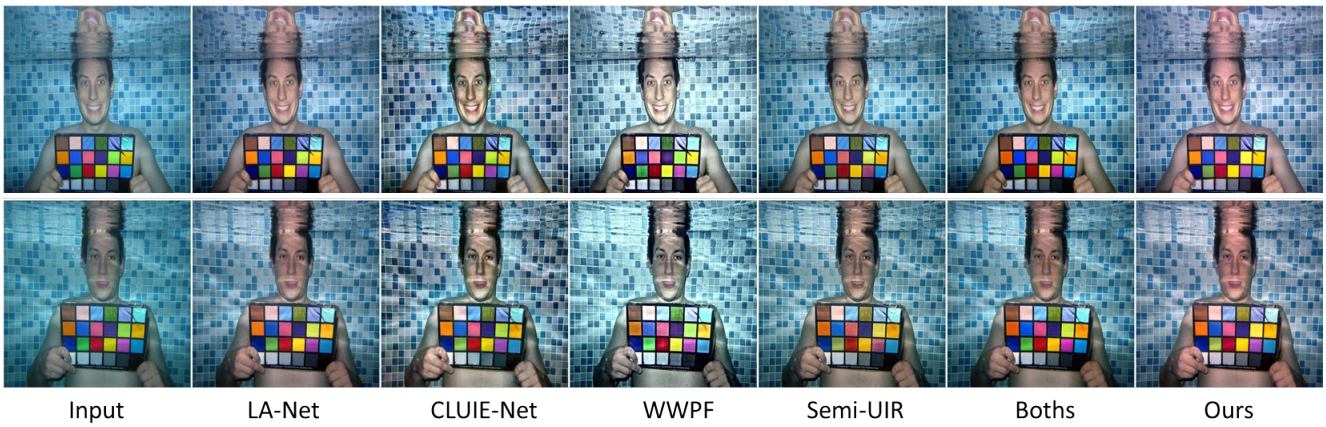

Figure S 10. Qualitative results comparison on Color-checker dataset with state-of-the-art (LANet [10], WaveNet [14], CLUIE [9], WWPF [16], Semi-UIE [4] and Boths [11]) methods for underwater image enhancement.

References

[1] Derya Akkaynak and Tali Treibitz. Sea-thru: A method for removing water from underwater images. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 1682–1691, 2019. 7

[2] Codruta O Ancuti, Cosmin Ancuti, Christophe De Vleeschouwer, and Philippe Bekaert. Color balance and fusion for underwater image enhancement. *IEEE Transactions on image processing*, 27(1):379–393, 2017. 7

[3] Sung-Jin Cho, Seo-Won Ji, Jun-Pyo Hong, Seung-Won Jung, and Sung-Jea Ko. Rethinking coarse-to-fine approach in single image deblurring. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 4641–4650, 2021. 3

[4] Shirui Huang, Keyan Wang, Huan Liu, Jun Chen, and Yunsong Li. Contrastive semi-supervised learning for underwater image restoration via reliable bank. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 18145–18155, 2023. 3, 5, 6, 7, 8, 9

[5] Md Jahidul Islam, Peigen Luo, and Junaed Sattar. Simultaneous enhancement and super-resolution of underwater imagery for improved visual perception. *arXiv preprint arXiv:2002.01155*, 2020. 5, 7

[6] Md Jahidul Islam, Youya Xia, and Junaed Sattar. Fast underwater image enhancement for improved visual perception. *IEEE Robotics and Automation Letters*, 5(2):3227–3234, 2020. 6, 7, 9

[7] Chongyi Li, Chunle Guo, Wenqi Ren, Runmin Cong, Junhui Hou, Sam Kwong, and Dacheng Tao. An underwater image enhancement benchmark dataset and beyond. *IEEE Transactions on Image Processing*, 29:4376–4389, 2019. 4, 6, 7

[8] Hanyu Li, Jingjing Li, and Wei Wang. A fusion adversarial underwater image enhancement network with a public test dataset. *arXiv preprint arXiv:1906.06819*, 2019. 5, 7

[9] Kunqian Li, Li Wu, Qi Qi, Wenjie Liu, Xiang Gao, Liqin Zhou, and Dalei Song. Beyond single reference for training: underwater image enhancement via comparative learning. *IEEE Transactions on Circuits and Systems for Video Technology*, 2022. 3, 5, 6, 7, 8, 9

[10] Shiben Liu, Huijie Fan, Sen Lin, Qiang Wang, Naida Ding, and Yandong Tang. Adaptive learning attention network for underwater image enhancement. *IEEE Robotics and Automation Letters*, 7(2):5326–5333, 2022. 5, 6, 7, 8, 9

[11] Xu Liu, Sen Lin, Kaichen Chi, Zhiyong Tao, and Yang Zhao. Boths: Super lightweight network-enabled underwater image enhancement. *IEEE Geoscience and Remote Sensing Letters*, 20:1–5, 2022. 3, 5, 6, 7, 8, 9

[12] Anish Mittal, Rajiv Soundararajan, and Alan C Bovik. Making a "completely blind" image quality analyzer. *IEEE Signal processing letters*, 20(3):209–212, 2012. 5, 6

[13] Karen Panetta, Chen Gao, and Sos Agaian. Human-visual-system-inspired underwater image quality measures. *IEEE Journal of Oceanic Engineering*, 41(3):541–551, 2015. 5, 6

[14] 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. 5, 6, 7, 8, 9

[15] Miao Yang and Arcot Sowmya. An underwater color image quality evaluation metric. *IEEE Transactions on Image Processing*, 24(12):6062–6071, 2015. 5, 6

[16] Weidong Zhang, Ling Zhou, Peixian Zhuang, Guohou Li, Xipeng Pan, Wenyi Zhao, and Chongyi Li. Underwater image enhancement via weighted wavelet visual perception fusion. *IEEE Transactions on Circuits and Systems for Video Technology*, 2023. 3, 5, 6, 7, 8, 9