Five A+ Network: You Only Need 9K Parameters for Underwater Image Enhancement

Jingxia Jiang[1*] Tian Ye[2*] Jinbin Bai[3*] Sixiang Chen[2] Wenhao Chai[4] Shi Jun[5] Yun Liu[6] Erkang Chen[1,7] ✉

School of Ocean Information Engineering, Jimei University [2]Hong Kong University of Science and Technology (GZ) [3]Department of Computer Science, National University of Singapore [4]University of Washington [5]School of Information Science and Engineering, Xinjiang University [6]College of Artificial Intelligence, Southwest University [7]Fujian Provincial Key Laboratory of Oceanic Information Perception and Intelligent Processing

Generalization Problem of Underwater Image Enhancement Networks

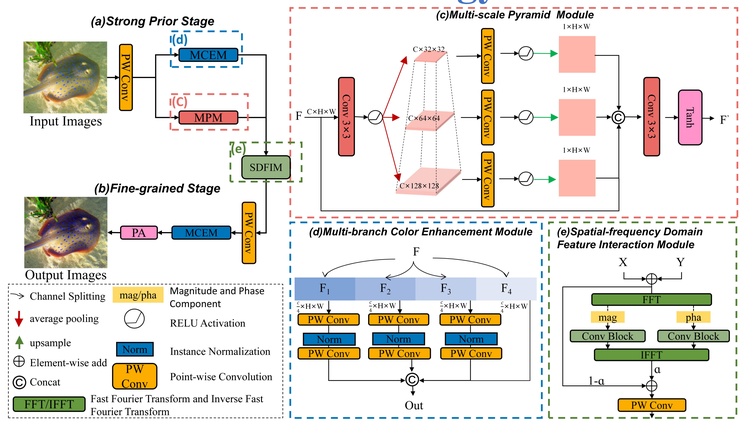
Insufficient computing and storage resources

- Insufficient computing and storage resources: traditional large-scale models struggle to achieve efficient underwater image enhancement on underwater robots and similar devices
- Recent **DNNs** that are speed-limited by high-resolution datasets.

Motivation

- Due to the limited resources of underwater robots and similar devices, traditional large-scale models struggle to achieve efficient underwater image enhancement on these platforms. Therefore, how to utilize deep learning algorithms for underwater image processing while meeting the real-time requirements is worth exploring.
- Furthermore, the underlying task of underwater image enhancement lacks large-scale high-quality datasets, and the performance of such real-world applications is largely jeopardized by data scarcity and various data degradations.
- So We propose a two-stage architecture that provides novel designs and directions for image enhancement. The strong prior stage decomposes mixed degradation into sub-problems, while the fine-grained stage focuses on enhancing the network's perception of intricate details.

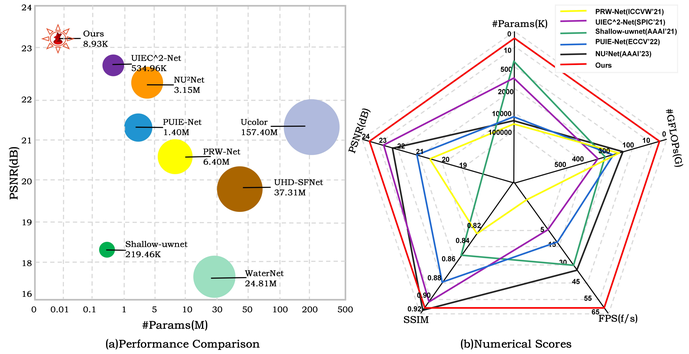
Methodology



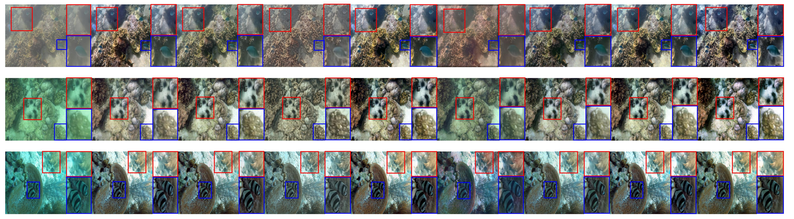
We introduce FA+Net, that reduces the the number of parameters of an enhancement model to 8.9K, which is approximately 10−100× fewer than previous methods.

FA+Net is the only model capable of real-time enhancement for 1080P images, efficiently running on an RTX 3090 GPU. It demonstrates exceptional performance across multiple datasets, making it a viable choice for deployment on mobile platforms.

The Computational Efficiency & Numerical Scores



Results and Analysis



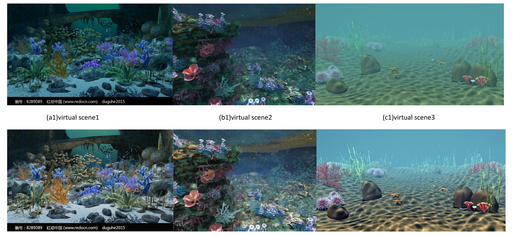
Visual comparisons with the SOTA methods on the T90.

Methods	T90 PSNR↑	T90 SSIM↑	T90 MSE↓	T90 UCIQE↑	T90 UIQM↑	C60 UCIQE↑	C60 UIQM↑	U45 UCIQE↑	U45 UIQM↑	Efficiency GFLOPs(G)↓	Efficiency #Params(M)↓	Efficiency #Runtime(s)↓	Efficiency FPS(f/s)↑
UDCP(ICCVW'13)[9]	13.415	0.749	0.228	0.572	2.755	0.560	1.859	0.574	2.275	-	-	42.13s	-
IBLA(TIP'17)[40]	18.054	0.808	0.142	0.582	2.557	0.584	1.662	0.565	2.387	-	-	-	-
WaterNet(TIP'19)[29]	16.305	0.797	0.161	0.564	2.916	0.550	2.113	0.576	2.957	193.70G	24.81M	0.680s	-
SMBL(TB'20)[46]	16.681	0.801	0.158	0.589	2.598	0.571	1.643	0.571	2.387	-	-	-	-
UWCNN(PR'20)[30]	17.949	0.847	0.221	0.517	3.011	0.492	2.222	0.527	3.063	-	-	-	-
PRW-Net(ICCVW'21)[18]	20.787	0.823	0.099	0.603	3.062	0.572	2.717	0.625	3.026	223.4G	6.30M	0.216s	4.624
Shallow-uwnet(AAAI'21)[37]	18.278	0.855	0.131	0.544	2.942	0.521	2.212	0.545	3.109	304.75G	0.22M	0.031s	31.836
UcolorTIP'21)[31]	21.093	0.872	0.096	0.555	3.049	0.530	2.167	0.554	3.148	443.85G	157.42M	2.758s	-
UIEC^2-Net(SPIC'21)[53]	22.958	0.907	0.078	0.599	2.999	0.580	2.228	0.604	3.125	367.53G	0.53M	0.174s	5.742
MLLE(TIP'22)[63]	19.561	0.845	0.115	0.592	2.624	0.581	1.977	0.597	2.454	-	-	-	-
UHD-SFNet(ACCV'22)[55]	18.877	0.810	0.144	0.559	2.551	0.528	1.741	0.585	2.826	15.24G	37.31M	0.059s	16.769
PUIE-Net(ECCV'22)[11]	21.382	0.882	0.093	0.566	3.021	0.543	2.155	0.563	3.192	423.05G	1.40M	0.071s	14.194
NU2Net(AAAI'23,Oral)[12]	22.419	0.923	0.086	0.587	2.936	0.555	2.222	0.593	3.185	146.64G	3.15M	0.024s	42.345
Ours	23.061	0.911	0.076	0.616	2.828	0.593	2.088	0.609	3.174	8.33G	0.009M	0.016s	60.724

The proposed RSFDM-Net achieves the best results on PSNR and UCIQE metrics which prove our proposed method is great at handling details textures and restore contrast. Also, FA+Net outperforms all other designs in terms of efficiency.

	256 × 256 patch #Flops(G)↓	#Runtime(s)↓	FPS(f/s)↑	512 × 512 patch #Flops(G)↓	#Runtime(s)↓	FPS(f/s)↑	1280 × 720 patch #Flops(G)↓	#Runtime(s)↓	FPS(f/s)↑	1920 × 1080 patch #Flops(G)↓	#Runtime(s)↓	FPS(f/s)↑	2560 × 1440 patch #Flops(G)↑	#Runtime(s)↓	FPS(f/s)↑
PRW-Net(ICCVW'21)[18]	15.836G	0.059s	16.799	63.34G	0.074s	13.401	223.40G	0.216s	4.624	Out of Memory			Out of Memory		
Shallow-uwnet(AAAI'21)[6]	21.67G	0.002s	396.948	86.69G	0.008s	115.600	304.75G	0.031s	31.836	685.70G	0.074s	13.418	1219.02G	0.129s	7.741
UIEC^2-Net(SPIC'21)[7]	26.14G	0.026s	37.739	104.54G	0.072s	13.722	367.53G	0.174s	5.742	826.93G	0.383s	2.607		Out of Memory	
PUIE-Net(ECCV'22)[1]	30.08G	0.009s	105.649	120.34G	0.020s	48.035	423.05G	0.071s	14.194	Out of Memory				Out of Memory	
NU2Net(AAAI'23,Oral)[2]	10.43G	0.002s	428.321	41.71G	0.007s	130.850	146.64G	0.024s	42.345	329.95G	0.051s	19.349	586.58G	0.090s	11.077
Ours	0.59G	0.003s	326.372	2.37G	0.007s	133.156	8.33G	0.016s	60.724	18.74G	0.033s	29.943	33.31G	0.057s	17.503

FA+Net exhibits a strong foundation for practical applications, as it presents a crushing victory for other methods on high-resolution images.



Also we have done some practice on the virtual scenes. Their colors become more brilliant.