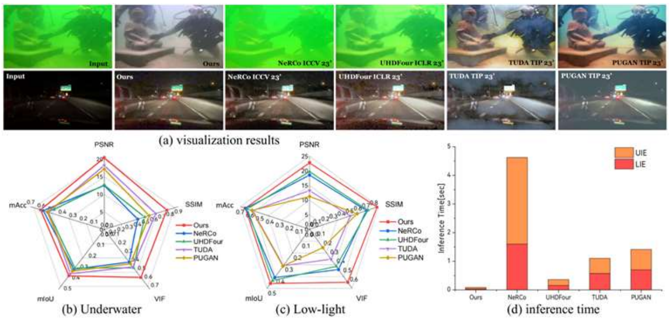
Retinex-Inspired Cooperative Game Through Multi-Level Feature Fusion for Robust, Universal Image Enhancement

BMVC 2024

Ruiqi Mao, Rongxin Cui*
r.cui@nwpu.edu.cn
School of Marine Science and Technology, Northwestern Polytechnical University

Motivation



Problem:

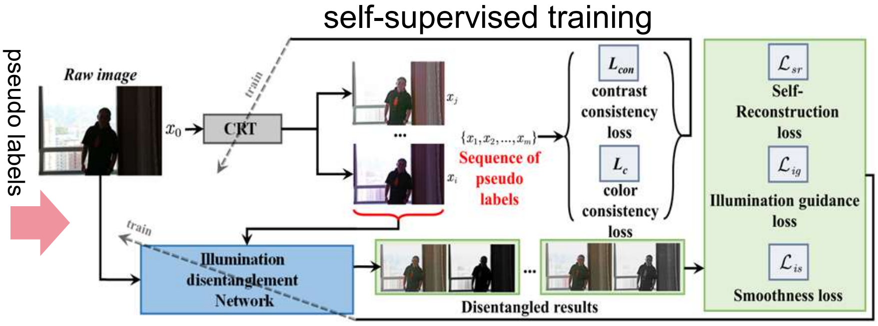
- Ineffectual in intricate scenarios(nighttime, underwater, backlit, etc.)
- Need paired dataset
- Hard to navigate performance and efficiency within a unified model

Methodology

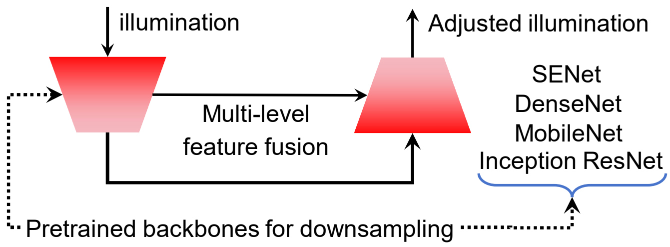
Phase I: self-supervised Retinex decomposition

Dynamic camera response Transformer

$$
\begin{cases}
x_t = f_{t-1}(x_{t-1}) \\
f_{t-1}: x_t = (1 - A_{t-1}) \sin\left(\frac{\pi}{2} x_{t-1}\right) + A_{t-1},\ A_{t-1} = \mathrm{CRT}(x_{t-1}) \\
\mathcal{R}_t, \mathcal{L}_t = \mathcal{K}(x_t), t = 1, 2, \cdots, m
\end{cases}
$$



Phase II: Flexible Illumination Adjustment



Phase III: Cooperative game

$$\min_{\alpha_f}\left\{\min_{\alpha_d,\omega}\mathcal{L}_{\text{game}}(\alpha_f,\alpha_d,\omega)\right\} \qquad \mathcal{L}_{\text{game}} := \mathcal{J}(\alpha_f) + \beta\mathbf{L}_D(\alpha_d,\omega)$$

Unsupervised loss for Flexible Illumination Adjustment

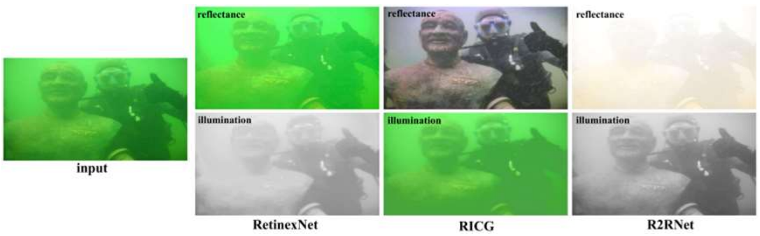
Self-supervised loss for Retinex decomposition

Users can choose different backbones according to navigate performance and efficiency

Experimental Results

Datasets	BAID			LSRW			UHD-LL			IT[sec]
Metrics	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	
URetinex-Net[24]	18.68	0.773	0.3081	14.78	0.661	0.4197	13.43	0.739	0.495	0.5785
UHDFour[2]	18.71	0.801	0.3176	18.20	0.656	0.3883	18.33	0.855	0.420	0.1088
LightenDiff [25]	19.88	0.855	0.3381	15.89	0.694	0.3563	16.23	0.789	0.447	0.0826
Wang et al. [26]	20.73	0.870	0.2682	16.05	0.706	0.3776	12.08	0.795	0.502	0.0028
GlobalDiff [27]	19.82	0.854	0.2986	13.82	0.685	0.3279	14.01	0.811	0.425	0.0976
CLIP-LIT [28]	22.35	0.862	0.3098	15.62	0.691	0.4087	13.12	0.651	0.470	0.1376
UNIE [29]	14.48	0.689	0.4628	10.35	0.562	0.4913	9.58	0.682	0.554	0.3675
NeRCo[5]	20.45	0.849	0.3281	14.20	0.653	0.4680	12.75	0.722	0.483	3.9918
Neural Preset[30]	18.05	0.726	0.3369	15.12	0.646	0.5091	12.36	0.708	0.582	0.0279
RICG	**22.65**	**0.886**	0.2927	**19.46**	**0.716**	0.3671	15.25	0.804	**0.413**	0.0306

Metrics / Methods	UCIQE↑	UIQM↑	UIQM: UICM	UIQM: UISM	UIQM: UICONM	CCF↑	CCF: Colorfulness	CCF: Contrast	CCF: Fogdensity
USUIR [11]	0.59	5.0744	2.78	7.39	0.787	29.2048	20.10	37.38	7.60
PUGAN [33]	0.58	4.7900	2.19	7.27	0.722	28.8607	14.97	38.02	**8.08**
Neural Preset [30]	0.54	5.0795	1.98	7.75	0.765	27.4928	14.01	36.53	7.26
TUDA [10]	**0.62**	5.0952	2.75	7.33	**0.798**	30.1717	**20.13**	38.86	7.74
Ours	0.60	**5.2013**	**2.86**	**7.80**	0.788	**30.8307**	18.21	**40.54**	7.62



Methods	**RICG**	NeRCo[5]	TUDA [10]	CLIP-LIT [28]	SCI [22]	UHDFour[2]	PUGAN[33]
mIoU	**0.4667**	0.3920	0.3728	0.3896	0.4343	0.4533	0.3804
mAcc	0.6067	0.5541	0.5446	0.5729	0.6011	**0.6093**	0.5259
aAcc	**0.7925**	0.7623	0.7015	0.7419	0.7535	0.7336	0.7126

More details about implementation code and visualization results of ablation can be accessed at *https://github.com/Ruiqi-Mao/RICG.*