Learning Semantic-Aware Knowledge Guidance for Low-Light Image Enhancement

Yuhui Wu[1], Chen Pan[1], Guoqing Wang[1]*, Yang Yang[1], Jiwei Wei[1], Chongyi Li[2], Heng Tao Shen[1]
[1]Center for Future Media, University of Electronic Science and Technology of China, China
[2]S-Lab, Nanyang Technological University, Singapore
wuyuhui132@gmail.com; panchen0103@163.com; gqwang0420@hotmail.com;
dlyyang@gmail.com; mathematic6@gmail.com; chongyi.li@ntu.edu.sg; shenhengtao@hotmail.com

CVPR 2023

Motivation

(a) Visual comparison on various scenes including car, human and sky.

Contributions

- We propose a semantic-aware knowledge-guided framework (SKF) to boost the performance of existing methods by jointly maintaining color consistency and improving image quality.

- We propose three key techniques to take full advantage of semantic priors provided by semantic knowledge bank (SKB): semantic-aware embedding (SE) module, semantic-guided color histogram (SCH) loss, and semantic-guided adversarial (SA) loss.

- We conduct experiments on LOL/LOL-v2 datasets and unpaired datasets. The experimental results demonstrate large performance improvements by our SKF, verifying its effectiveness in resolving the LLIE task.

Problem definition of semantic-aware LLIE

$$M = \mathbf{F}_{segment}(I_l; \theta_s), \qquad (1)$$

$$\widehat{I_h} = \mathbf{F}_{enhance}(I_l, M; \theta_e), \qquad (2)$$

$$\widehat{\theta_e} = argmin \mathcal{L}(\widehat{I_h}, I_h, M), \qquad (3)$$

Framework

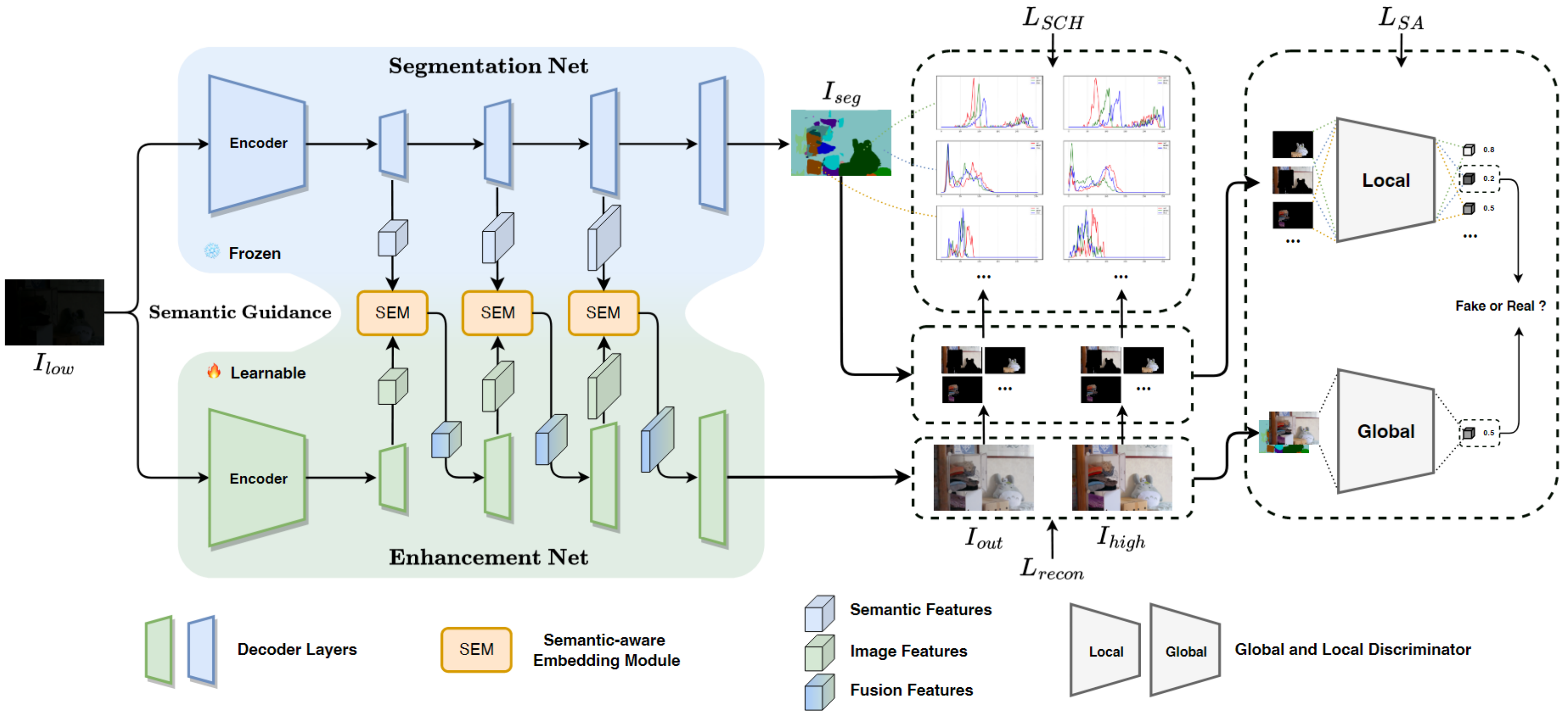


Figure 2. Overview of our Semantic-aware Knowledge-guided Framework (SKF). With a pre-trained Segmentation Net, our SKF utilizes semantic priors to improve the enhancement process in two aspects: **(a)** In feature-level, the multi-scale semantic-aware embedding modules enable cross-modal interactions between semantic features and image features in representation space. **(b)** In loss-level, the semantic segmentation result is introduced into the computation of color histogram loss and adversarial loss as a guidance.

SKF
- Feature-level:SE module
- Loss-level
  - SCH loss
  - SA loss

Semantic-Aware Embedding Module

$$A^b = Softmax\left(W_k(F_i^b) \times W_q(F_s^b)/\sqrt{C}\right), \qquad (4)$$

$$F_o^b = FN(W_v(F_i^b) \times A^b + F_i^b), \qquad (5)$$

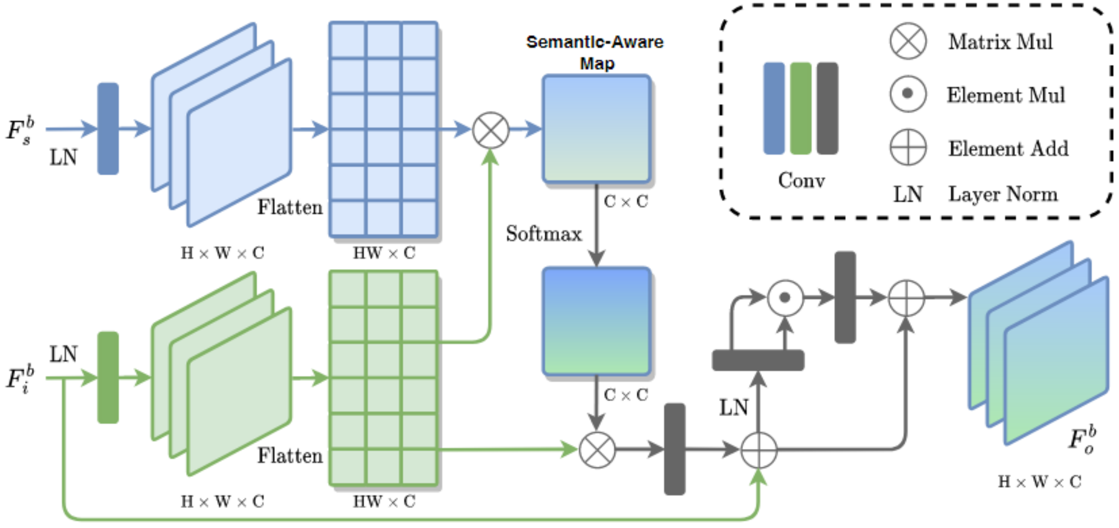
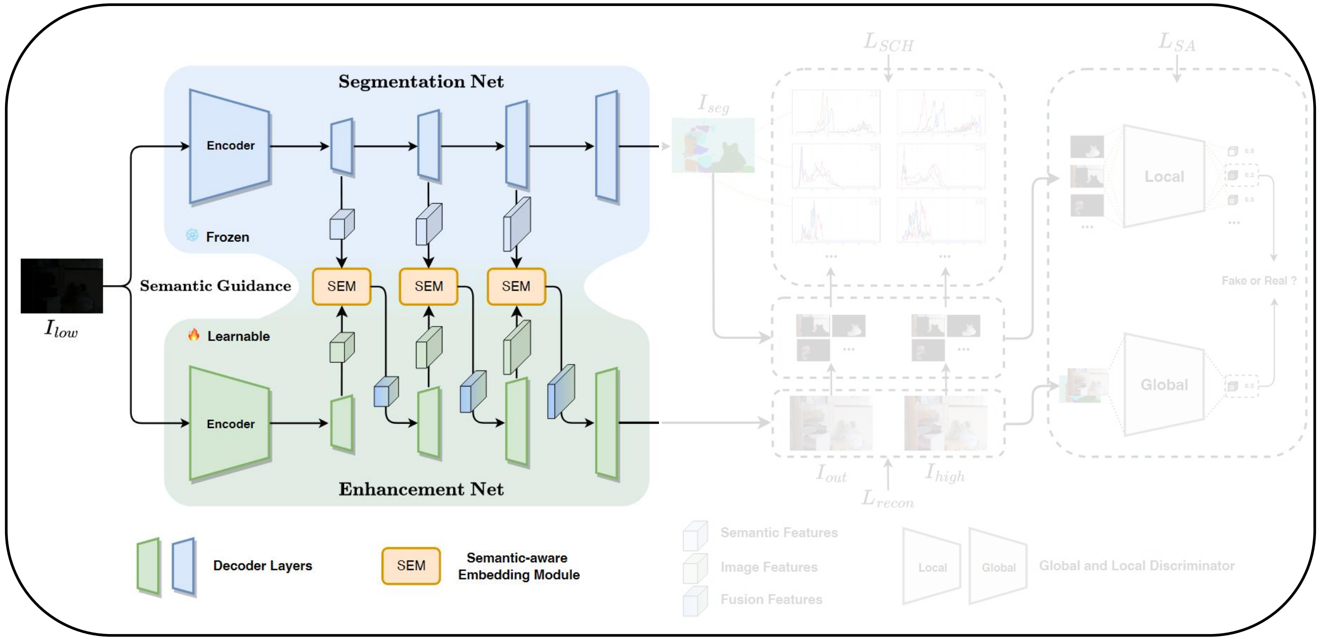


Figure 3. Architecture of the semantic-aware embedding (SE) module. At the b^{th} decoder layer, SE module transforms the image feature map F_i^b with the semantic feature map F_s^b and produces the refined output feature F_o^b.



Semantic-Guided Color Histogram Loss

$$P = \{P^0, P^1, \ldots, P^{class}\}, \quad P^c = I_{out} \odot I^c_{seg}, \tag{6}$$

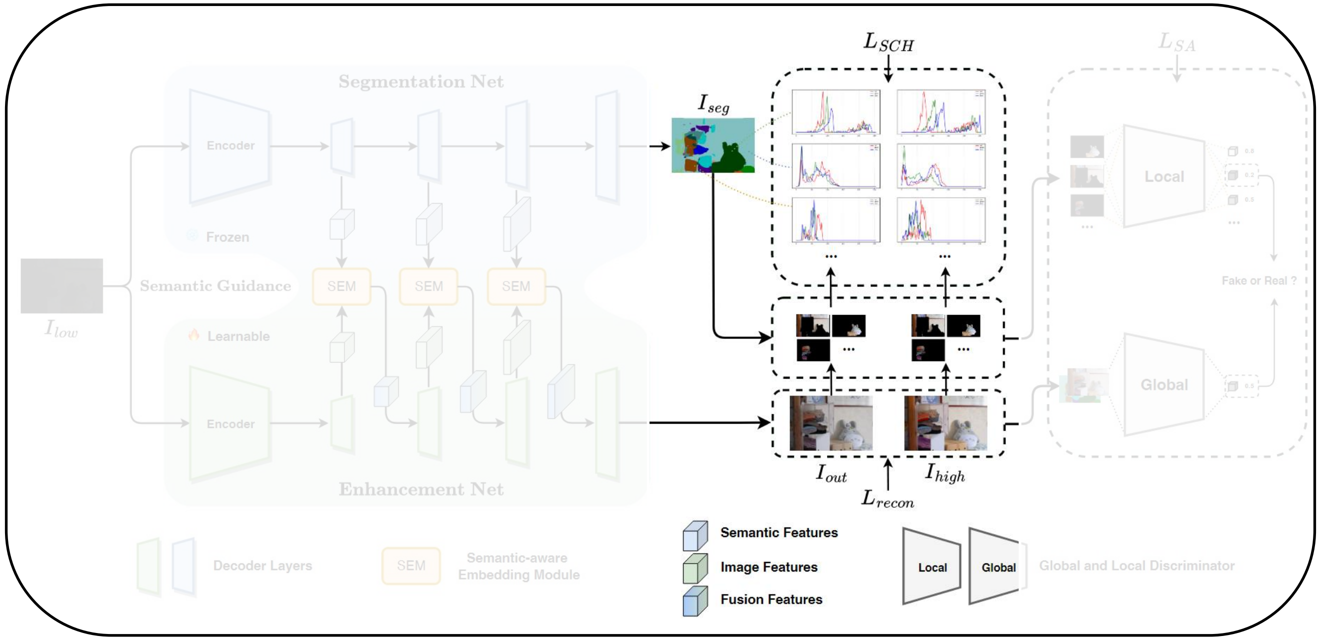
kernel density estimation

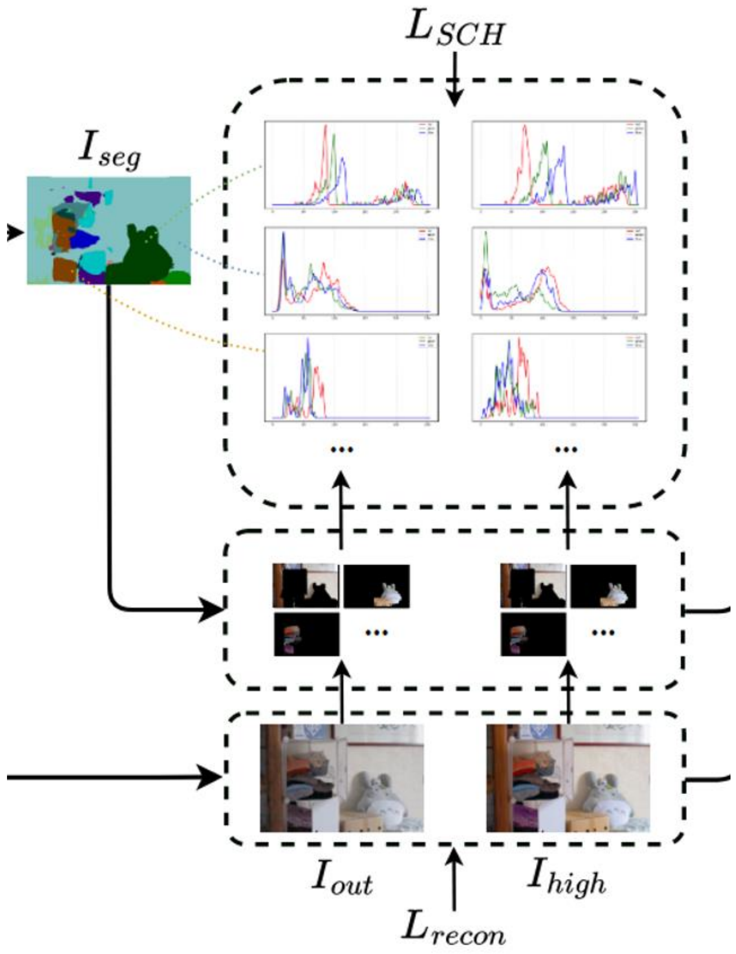
$$x^h_{ij} = x_j - \frac{i - 0.5}{255}, \quad x^l_{ij} = x_j - \frac{i + 0.5}{255}, \tag{7}$$

$$H^c_i = \sum_j \left(Sigmoid(\alpha \cdot x^h_{ij}) - Sigmoid(\alpha \cdot x^l_{ij}) \right), \tag{8}$$

$$H^c = \{i, H^c_i\}_{i=0}^{255}, \tag{9}$$

$$\mathcal{L}_{SCH} = \sum_c \| H^c(\hat{I}_h) - H^c(I_h) \|_1, \tag{10}$$





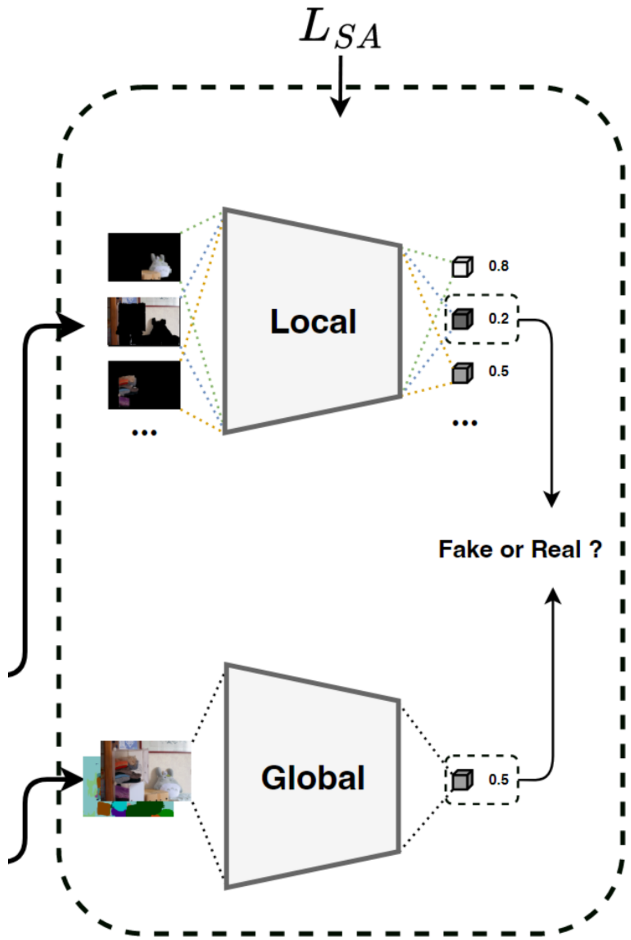
Semantic-Guided Adversarial Loss

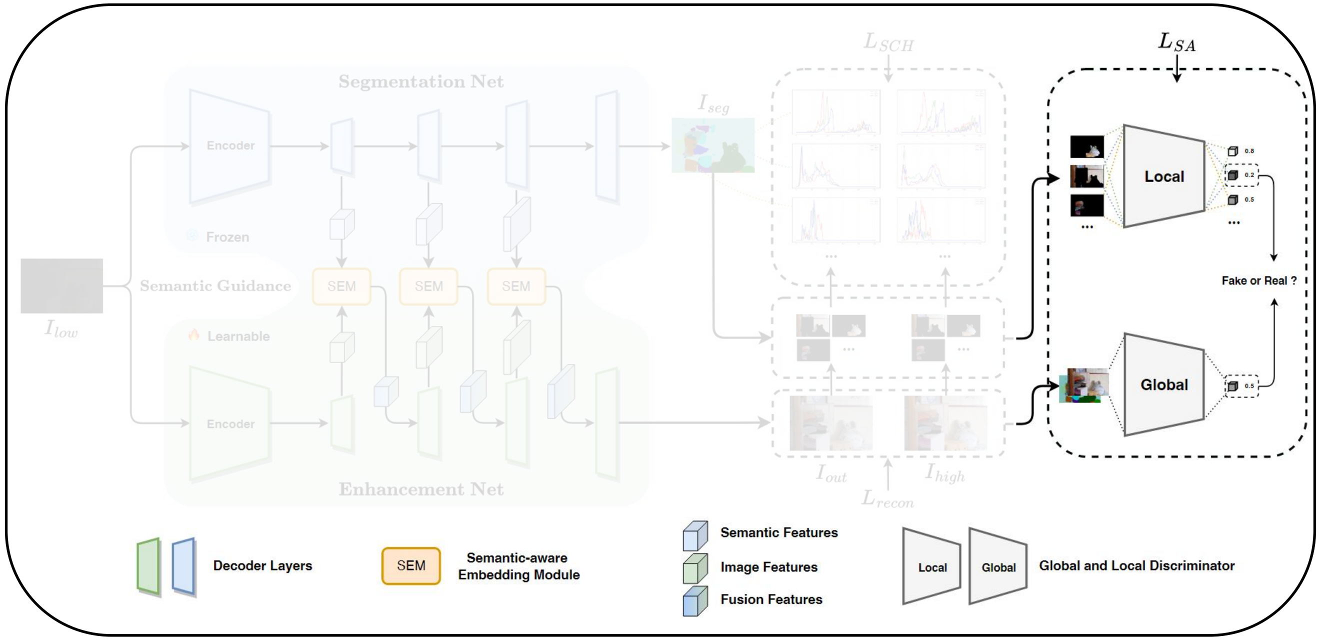
$$\mathcal{L}_{local} = \min_{G} \max_{D} \mathbb{E}_{x_r \sim p_{real}} MSE(D(x_r), 0) + \mathbb{E}_{x_f \sim p_{fake}} MSE(D(x_f), 1), \tag{11}$$

$$x_f = P^t, D(P^t) = \min(D(P^0), \ldots, D(P^{class})), \tag{12}$$

$$\mathcal{L}_{global} = \min_{G} \max_{D} \mathbb{E}_{x_r \sim p_{real}} MSE(D(x_r), 0) + \mathbb{E}_{x_f \sim p_{fake}} MSE(D(x_f, I'_{seg}), 1), \tag{13}$$

$$\mathcal{L}_{SA} = \mathcal{L}_{global} + \mathcal{L}_{local}, \tag{14}$$

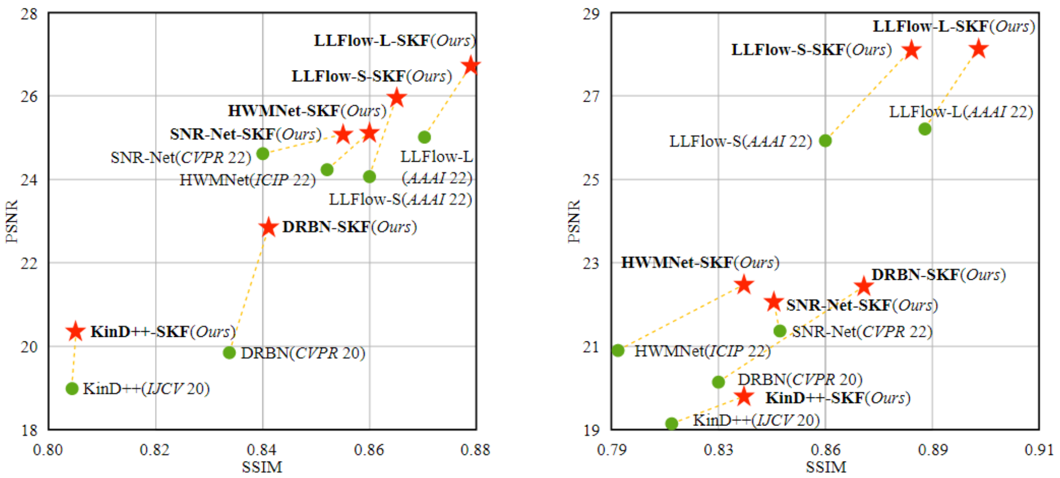




Experiments

Table 1. Quantitative comparison on the LOL [43] and LOL-v2 [49] datasets. ↑ (↓) denotes that, larger (smaller) values lead to better quality. + (-) denotes the improvement (reduction) of performance, corresponding to ↑ (↓). The bold denotes the best.

Method	LOL				LOL-v2				Param(M)
	PSNR ↑	SSIM ↑	LPIPS ↓	NIQE ↓	PSNR ↑	SSIM ↑	LPIPS ↓	NIQE ↓	
LIME [13] TIP'16	16.760	0.560	0.350	-	15.240	0.470	-	-	-
Zero-DCE [11] CVPR'20	14.861	0.562	0.335	7.767	18.059	0.580	0.313	8.058	0.33
EnlightGAN [16] TIP'21	17.483	0.652	0.322	4.684	18.640	0.677	0.309	5.089	8.64
ISSR [8] MM'20	18.846	0.788	0.243	5.249	16.994	0.798	0.206	5.179	12.12
MIRNet [51] PAMI'22	24.140	0.842	0.131	4.203	20.357	0.782	0.317	5.094	5.90
RetinexNet [43] BMVC'18	16.770	0.462	0.474	8.873	18.371	0.723	0.365	5.849	0.62
RetinexNet-SKF(Ours)	20.418 (+3.648)	0.711 (+0.249)	0.216 (+0.258)	4.211 (+4.662)	19.849 (+1.478)	0.719 (-0.004)	0.255 (+0.110)	4.233 (+1.616)	0.66
KinD [54] MM'19	20.870	0.799	0.207	5.189	17.544	0.669	0.375	6.849	8.03
KinD-SKF(Ours)	21.913 (+1.043)	0.835 (+0.036)	0.143 (+0.064)	5.031 (+0.158)	19.821 (+2.277)	0.833 (+0.164)	0.201 (+0.174)	4.778 (+2.071)	8.50
DRBN [48] CVPR'20	19.860	0.834	0.155	4.793	20.130	0.830	0.147	4.961	2.21
DRBN-SKF(Ours)	22.837 (+2.977)	0.841 (+0.007)	0.138 (+0.017)	4.464 (+0.329)	22.441 (+2.311)	0.871 (+0.041)	0.132 (+0.015)	4.460 (+0.501)	2.43
KinD++ [53] IJCV'20	18.970	0.804	0.175	4.760	19.087	0.817	0.180	5.086	9.63
KinD++-SKF(Ours)	20.363 (+1.393)	0.805 (+0.001)	0.201 (-0.026)	**4.142** (+0.618)	19.779 (+0.692)	0.837 (+0.020)	0.178 (+0.002)	4.179 (+0.907)	10.21
HWMNet [7] ICIP'22	24.240	0.852	0.114	5.141	20.928	0.798	0.359	5.970	66.56
HWMNet-SKF(Ours)	25.086 (+0.846)	0.860 (+0.008)	0.108 (+0.006)	4.346 (+0.795)	22.490 (+1.562)	0.836 (+0.038)	0.175 (+0.184)	4.683 (+1.288)	69.98
SNR-LLIE-Net [46] CVPR'22	24.608	0.840	0.151	5.179	21.479	0.848	0.157	4.623	39.13
SNR-LLIE-Net-SKF(Ours)	25.031 (+0.552)	0.855 (+0.015)	0.113 (+0.038)	4.722 (+0.457)	21.927 (+0.448)	0.842 (-0.006)	0.160 (-0.003)	**3.963** (+0.660)	39.44
LLFlow-S [41] AAAI'22	24.060	0.860	0.136	5.412	25.922	0.860	0.173	6.150	4.97
LLFlow-S-SKF(Ours)	25.942 (+1.882)	0.865 (+0.005)	0.125 (+0.011)	5.606 (-0.194)	28.107 (+2.185)	0.884 (+0.024)	0.133 (+0.040)	5.415 (+0.735)	5.26
LLFlow-L [41] AAAI'22	24.999	0.870	0.117	5.582	26.200	0.888	0.137	5.406	37.68
LLFlow-L-SKF(Ours)	**26.798** (+1.799)	**0.879** (+0.009)	**0.105** (+0.012)	5.589 (-0.007)	**28.451** (+2.251)	**0.905** (+0.017)	**0.112** (+0.025)	5.725 (-0.319)	39.91



(b) Performance comparison on LOL/LOL-v2 (left/right) datasets

Table 2. Quantitative comparison on the LOL [43], LOL-v2 [49], MEF [31], LIME [12], NPE [39] and DICM [21] datasets in terms of NIQE, where smaller values lead to better quality.

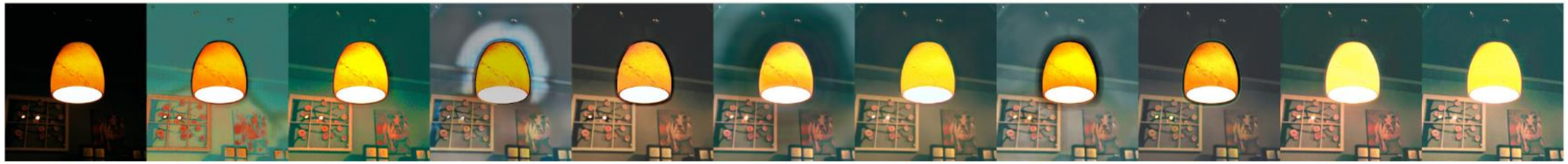
Method	LOL	LOL-v2	MEF	LIME	NPE	DICM
Input	6.7488	6.7911	4.2650	4.4380	4.3124	4.2550
RetinexNet [43] BMVC'18	6.8731	5.8488	4.1490	4.4200	4.5008	4.5912
RetinexNet-SKF(Ours)	4.2118	4.2331	**3.6321**	4.0779	4.0152	3.6945
KinD [54] MM'19	5.1891	6.8490	4.1344	4.6418	4.6896	3.9371
KinD-SKF(Ours)	5.0306	4.7783	3.9460	4.3607	3.8721	3.7909
DRBN [48] CVPR'20	4.7930	4.9612	4.0956	4.4019	3.9205	4.0433
DRBN-SKF(Ours)	4.4636	4.4599	4.0894	4.3392	4.0192	3.8541
KinD++ [53] IJCV'20	4.7602	5.0856	3.7498	4.3756	3.9848	3.7076
KinD++-SKF(Ours)	**4.1415**	**4.1785**	3.7645	**3.9892**	**3.8201**	**3.5382**
HWMNet [7] ICIP'22	5.1407	5.9702	4.2175	4.3549	4.0683	3.9196
HWMNet-SKF(Ours)	4.3460	4.6826	4.0312	4.3699	3.9942	4.0760

Experiments

(a) Input (b) GT (c) RetinexNet (d) RetinexNet-SKF (e) KinD (f) KinD-SKF (g) DRBN (h) DRBN-SKF

(i) KinD++ (j) KinD++-SKF (k) HWMNet (l) HWMNet-SKF (m) SNR-LLIE-Net (n) SNR-LLIE-Net-SKF (o) LLFlow-L (p) LLFlow-L-SKF

Figure 4. Visual comparison of baseline methods with and without SKF on LOL dataset. Our SKF enables baseline methods produce images with less noise, more color information and realistic details.



(a) Input (b) RetinexNet (c) RetinexNet-SKF (d) KinD (e) KinD-SKF (f) DRBN (g) DRBN-SKF (h) KinD++ (i) KinD++-SKF (j) HWMNet (k) HWMNet-SKF

Figure 5. Visual comparison of baseline methods with and without SKF on LIME dataset.

Experiments

(a) Baseline (b) W/ SCH loss (c) W/ SE (d) W/SCH loss & SE (e) W/SCH loss & SA loss & SE (f) GT

Figure 6. Visual comparison of DRBN-SKF for investigating the contribution of key techniques of our SKF.

Table 3. Ablation study of KinD++-SKF, DRBN-SKF and HWMNet-SKF for investigating the contribution of key techniques of our SKF.

SCH loss	SA loss	SE module	KinD++-SKF				DRBN-SKF				HWMNet-SKF			
			PSNR ↑	SSIM ↑	LPIPS ↓	NIQE ↓	PSNR ↑	SSIM ↑	LPIPS ↓	NIQE ↓	PSNR ↑	SSIM ↑	LPIPS ↓	NIQE ↓
			18.970	0.804	0.175	4.760	19.860	0.834	0.155	4.793	24.240	0.852	0.114	5.141
✓			19.170	0.806	**0.170**	4.759	20.040	0.835	0.154	4.793	24.590	0.859	0.112	5.023
	✓		19.385	0.800	0.189	4.392	20.070	0.834	0.149	4.701	24.305	0.853	0.111	4.712
		✓	19.781	0.808	0.181	4.712	21.334	0.837	0.143	4.678	24.477	0.859	0.111	4.988
✓		✓	**20.620**	**0.815**	0.176	4.536	22.550	0.836	0.150	4.581	**25.123**	0.860	0.111	4.711
✓	✓	✓	20.363	0.805	0.201	**4.142**	**22.837**	**0.841**	**0.138**	**4.464**	25.086	**0.860**	**0.108**	**4.346**

Experiments

Table 4. Ablation study of HWMNet-SKF for investigating the effect of semantic priors in the loss function.

$\mathcal{L}_{SCH}$ w/o S	$\mathcal{L}_{SCH}$ w/ S	$\mathcal{L}_{SA}$ w/o SA	$\mathcal{L}_{SA}$ w/o S	$\mathcal{L}_{SA}$ w/ S	**PSNR ↑**	**SSIM ↑**	**LPIPS ↓**	**NIQE ↓**
					24.240	0.852	0.114	5.141
✓		✓			24.477	0.859	0.111	4.988
✓			✓		24.568	0.857	0.113	4.613
✓				✓	24.668	0.859	0.111	4.567
	✓	✓			**25.123**	**0.860**	**0.108**	4.711
	✓		✓		25.040	0.859	0.111	4.546
	✓			✓	25.086	**0.860**	**0.108**	**4.311**

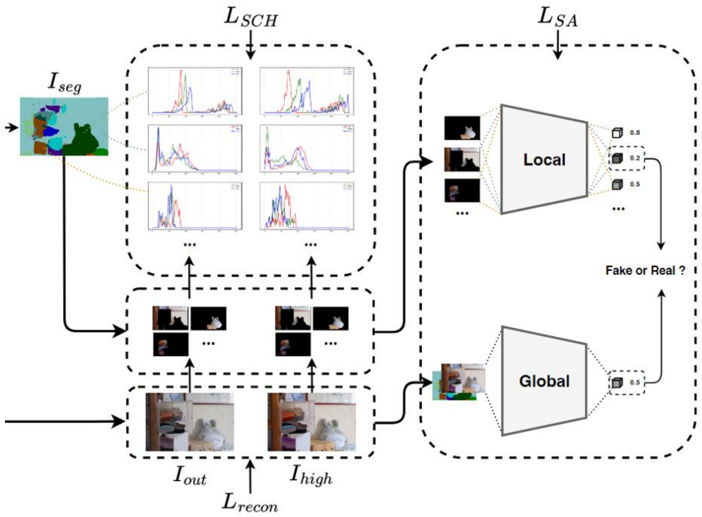


Table 5. Ablation study for investigating whether the performance improvement comes from semantic priors or more parameters.

Method		PSNR ↑	SSIM ↑	LPIPS ↓	Param(M)
HWMNet	Baseline	24.240	0.852	0.114	66.56
	Large	24.445	0.853	0.115	69.99
	w/ SKF	**25.086**	**0.860**	**0.108**	69.98
LLFlow-S	Baseline	24.060	0.860	0.136	4.97
	Large	24.167	0.858	0.137	5.38
	w/ SKF	**25.942**	**0.865**	**0.125**	5.26
LLFlow-L	Baseline	24.999	0.870	0.117	37.68
	Large	25.292	0.873	0.113	40.55
	w/ SKF	**26.798**	**0.879**	**0.105**	39.91

Thanks!