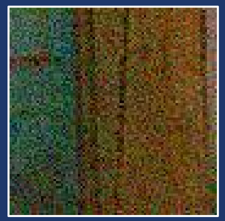
HVI: A New Color Space for Low-light Image Enhancement

Qingsen Yan[1*] Yixu Feng[1*] Cheng Zhang[1] Guansong Pang[2]
Kangbiao Shi[1] Peng Wu[1] Wei Dong[3] Jinqiu Sun[1] Yanning Zhang[1]

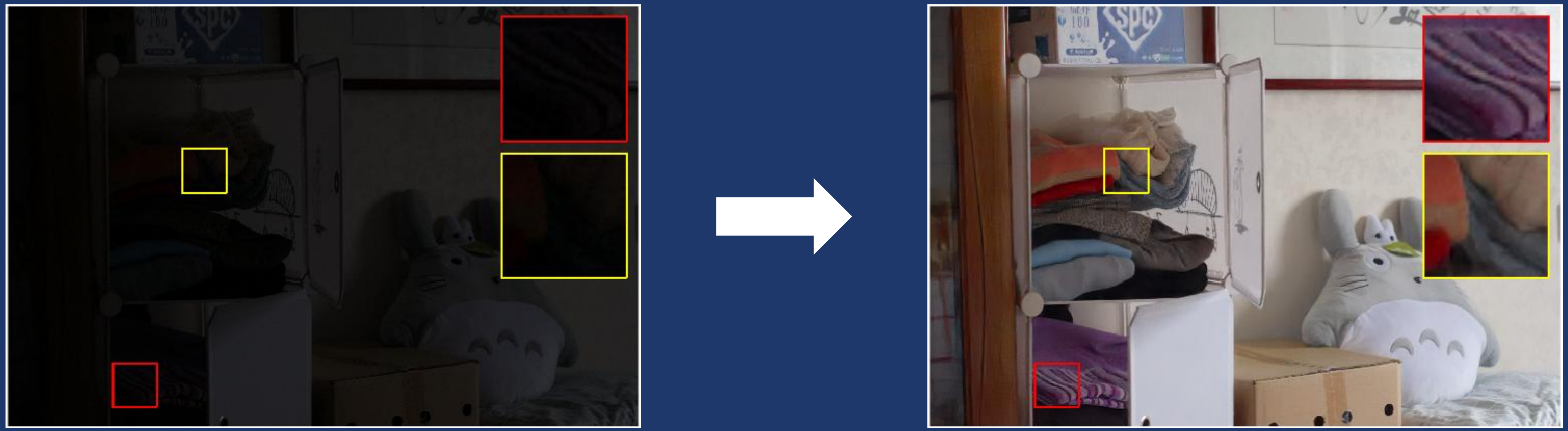
[1]Northwestern Polytechnical University
[2]Singapore Management University
[3]Xi'an University of Architecture and Technology

Low Light Image Enhancement

- Low-light Conditions: severe noise, poor visual quality, color distortion.
- Low-Light Image Enhancement (LLIE):
- 1. Improving the image brightness;
- 2. Reducing the impact of noise and color bias.

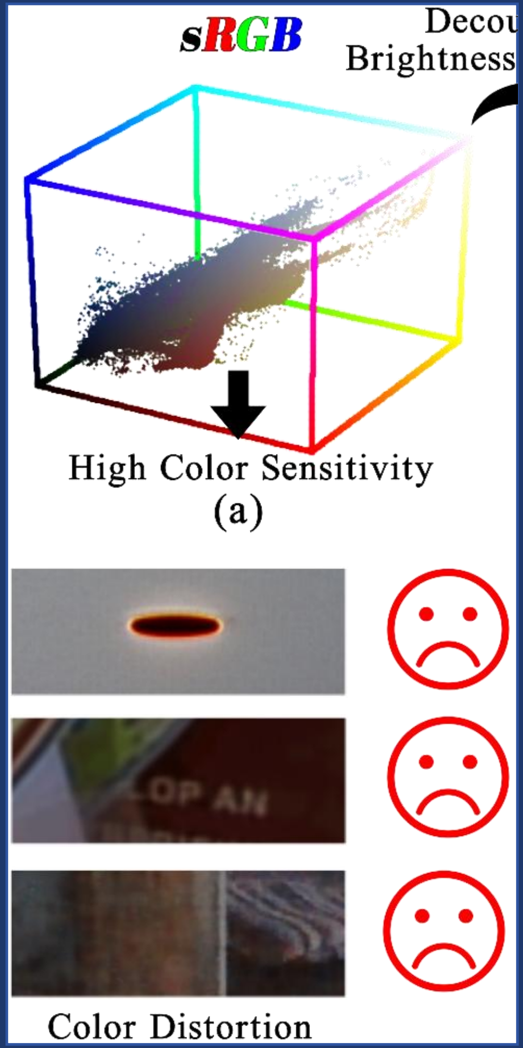


Noise Visualization



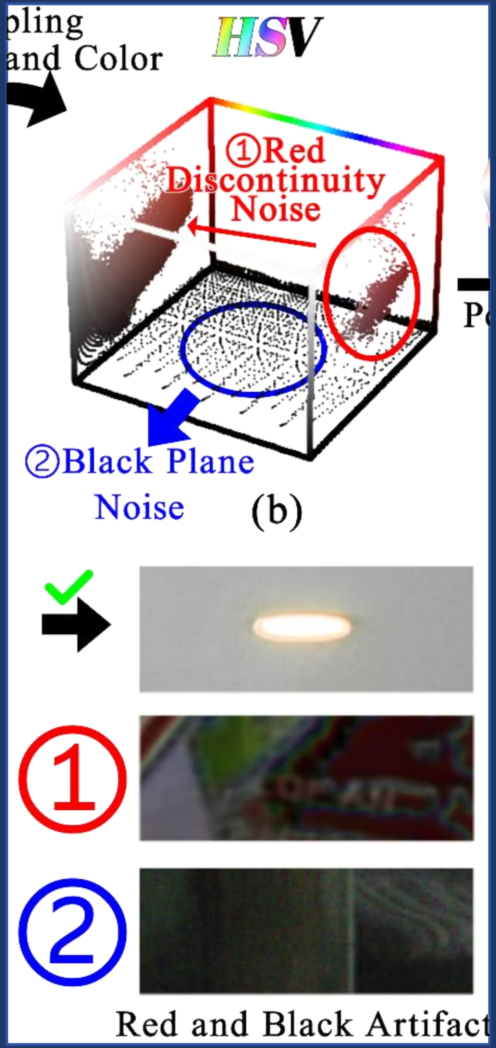
Motivation: sRGB Sensitivity

- The majority of existing LLIE method:
- 1. Employing deep neural networks to learn a mapping relationship between low-light images and normal-light images within the standard RGB (sRGB) space;
- 2. Weakness: brightness is coupling with the color from the three sRGB channels, *a.k.a* high color sensitivity, causing an obvious color distortion of the restored image in these LLIE methods.

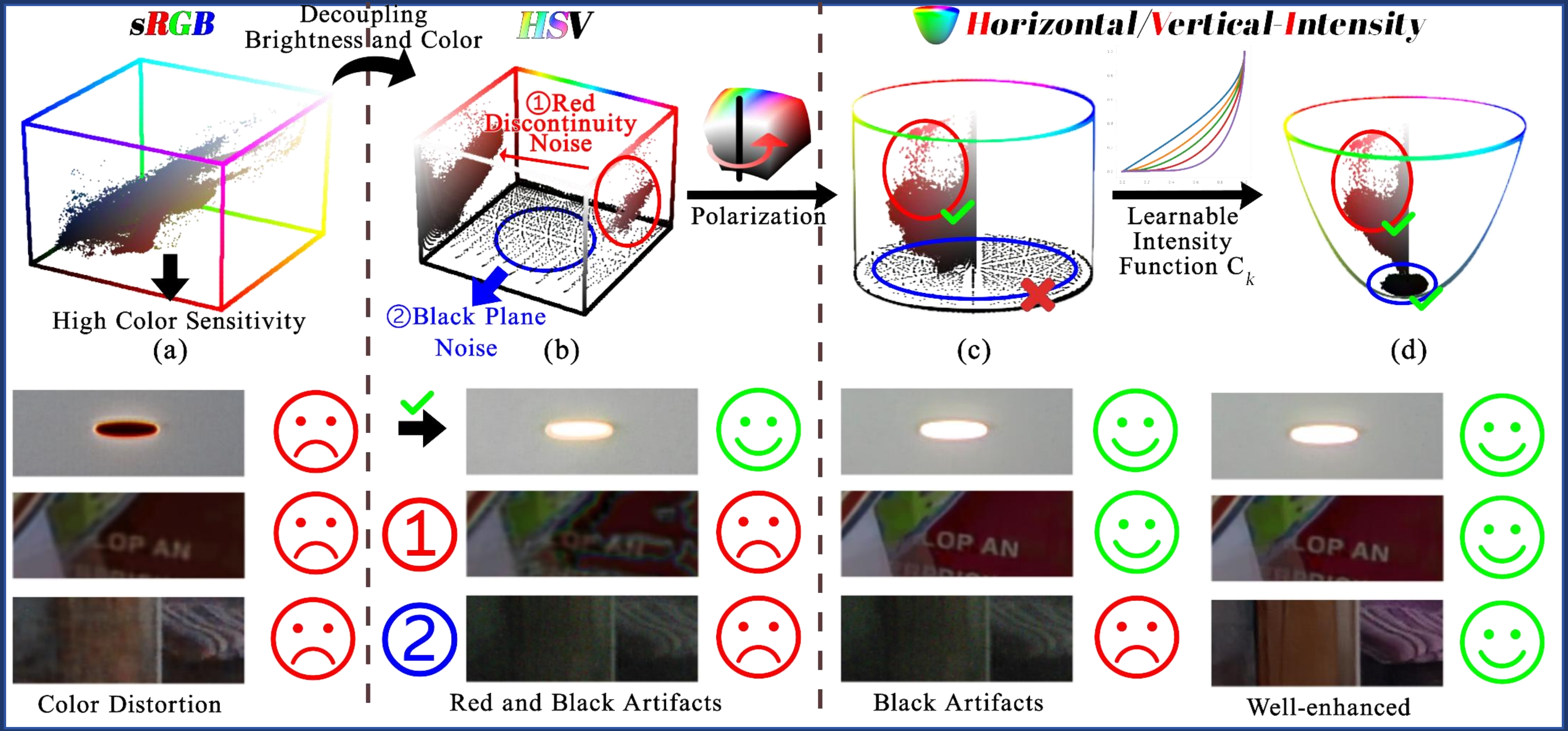


Motivation: Noise in HSV

- Recent methods have sought to transform images from the sRGB color space to the Hue, Saturation and Value (HSV) color space.
- Two types of HSV space noise:
- ① Red Discontinuity Noise
- ② Black Plane Noise
- Resulting in increased Euclidean distances for similar color and the introduction of artifacts in the final images.



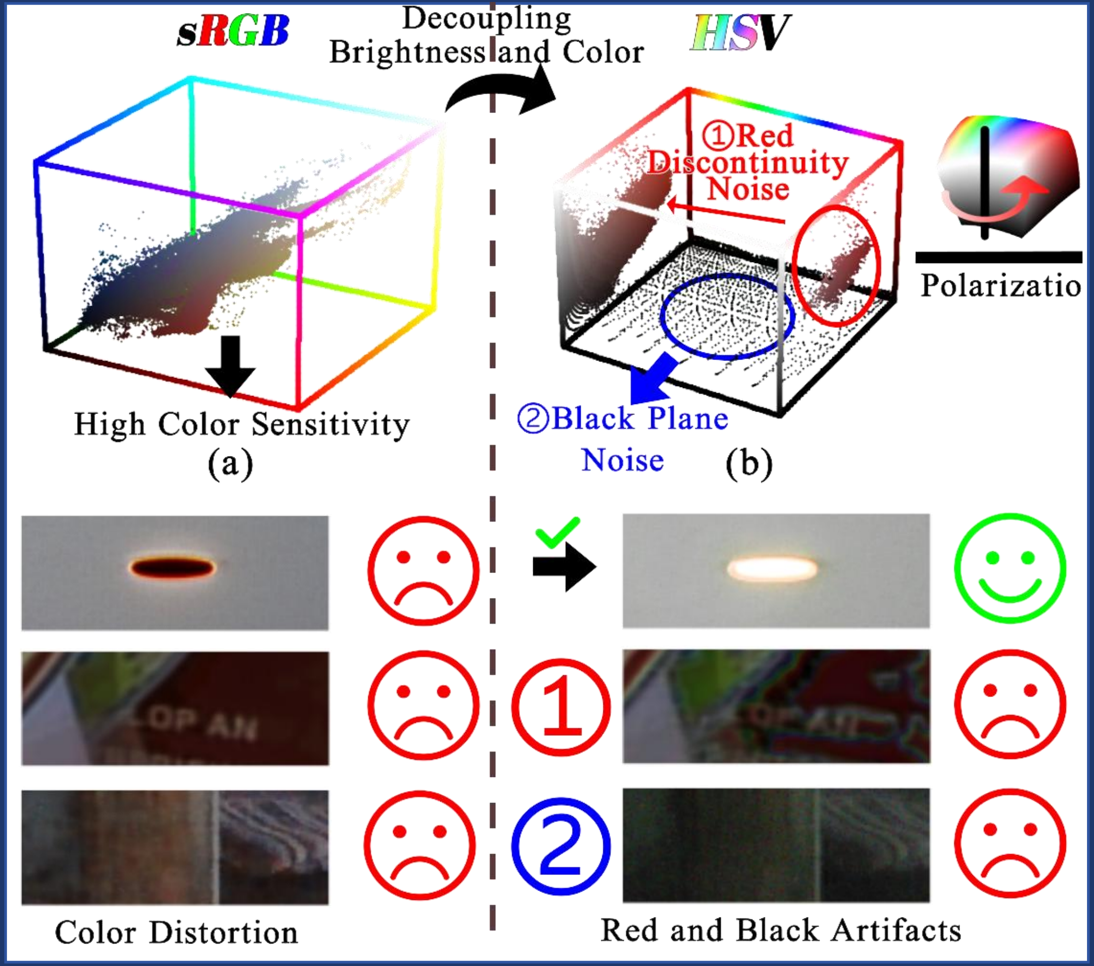
Horizontal/Vertical-Intensity



- Therefore, we propose a new color space： Horizontal/Vertical-Intensity (HVI), to decouple brightness and color, and solve two types of noise in HSV.
- The key intuition is that minimizing color space noise, by reducing Euclidean distances in similar colors.

HVI Transformation

Step One：sRGB → HSV

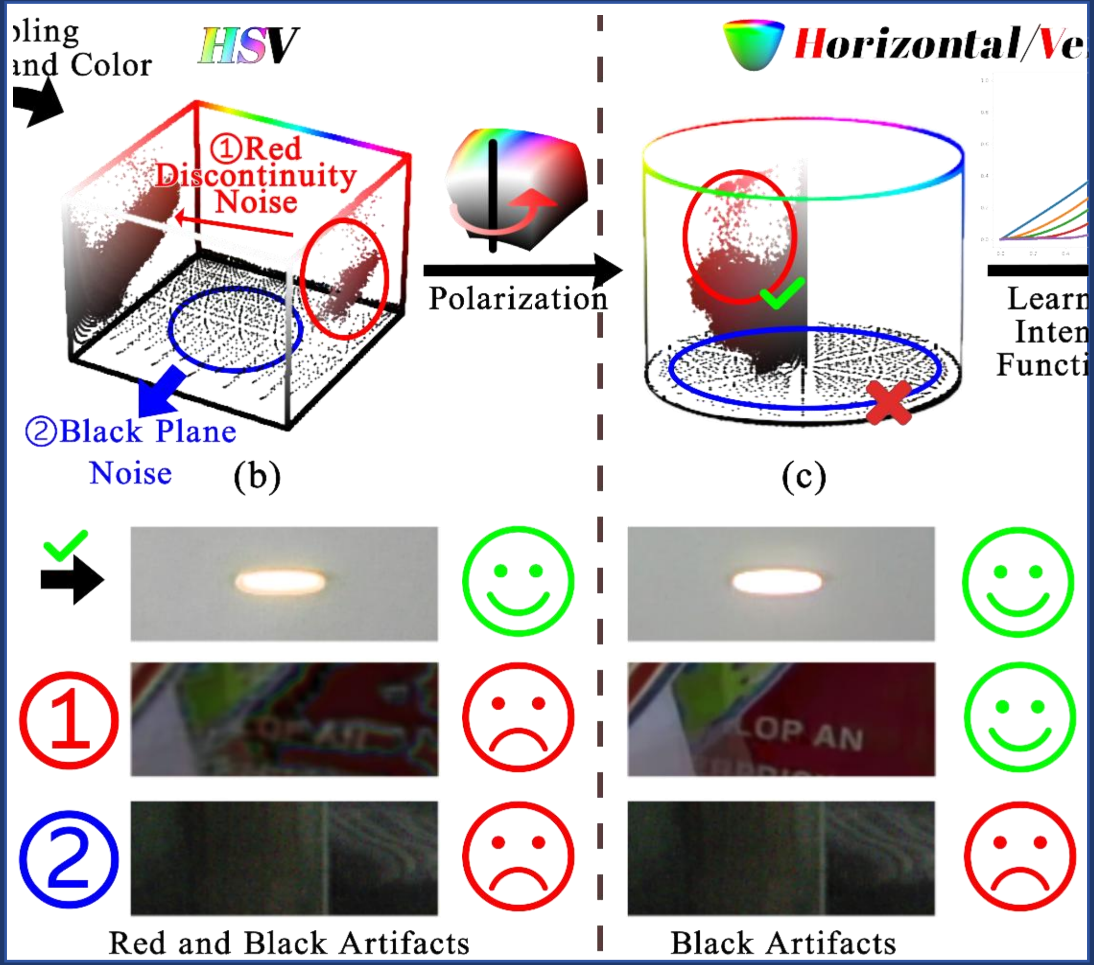


- 1. Estimating the illumination intensity map of the scene from a sRGB input image： $\mathbf{I}_{max}(x) = \max_{c\in\{R,G,B\}} (\mathbf{I}_c(x)).$
- 2. Decoupling Brightness (I_{max}) and Color (Hue-Saturation Plane):
- (1) Saturation: $\mathbf{s} = \begin{cases} 0, & \mathbf{I}_{max} = 0 \\ \frac{\Delta}{\mathbf{I}_{max}}, & \mathbf{I}_{max} \neq 0 \end{cases},$
- (2) Hue: $\mathbf{h} = \begin{cases} 0, & \text{if } \mathbf{s} = 0 \\ \frac{\mathbf{I_G} - \mathbf{I_B}}{\Delta} \mod 6, & \text{if } \mathbf{I}_{max} = \mathbf{I_R} \\ 2 + \frac{\mathbf{I_B} - \mathbf{I_R}}{\Delta}, & \text{if } \mathbf{I}_{max} = \mathbf{I_G} \\ 4 + \frac{\mathbf{I_R} - \mathbf{I_G}}{\Delta}, & \text{if } \mathbf{I}_{max} = \mathbf{I_B} \end{cases},$

$$\Delta = \mathbf{I}_{max} - min(\mathbf{I}_c)$$

HVI Transformation

Step Two: Polarization



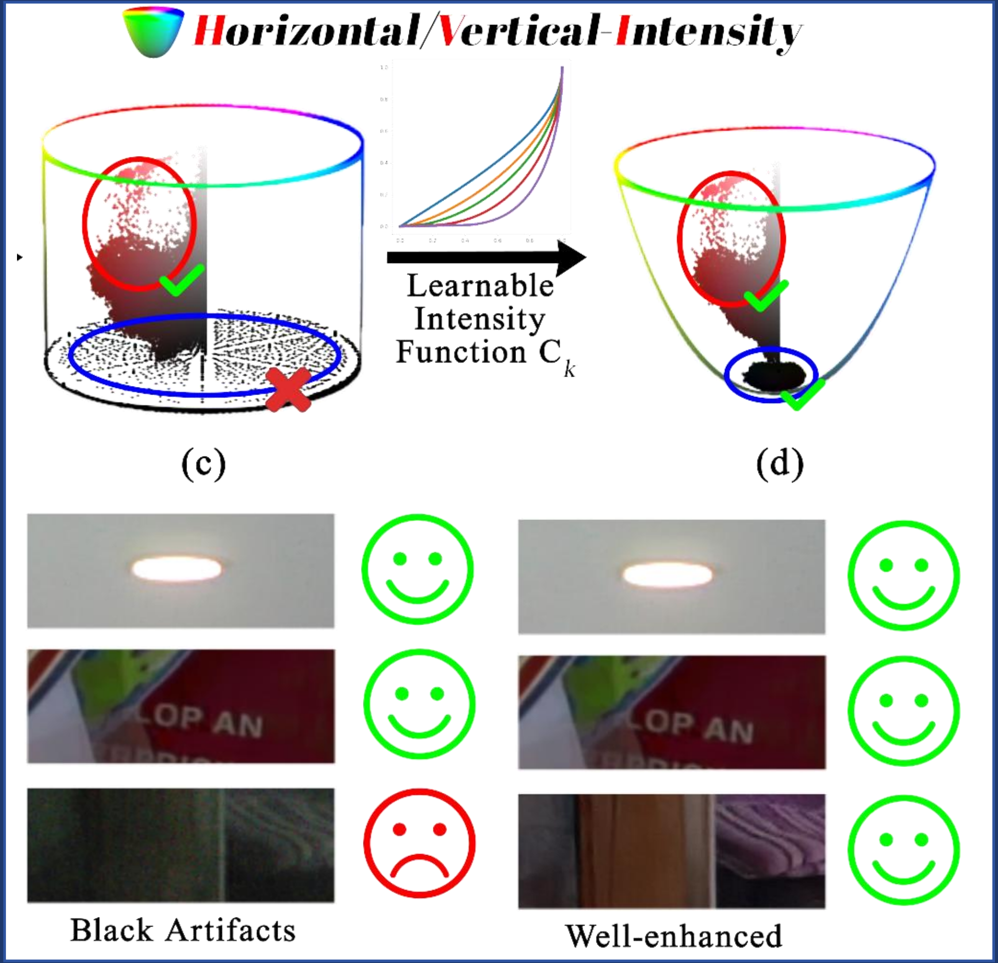
- Reduce ① Red Discontinuity Noise
- Along the Hue axis, the color red appears identically at both h = 0 and h = 6, which splits the same color across two ends of the spectrum.
- Polarize the Hue (**h**) Channel:

$$h = \cos\left(\frac{\pi \mathbf{h}}{3}\right), \text{and } v = \sin\left(\frac{\pi \mathbf{h}}{3}\right).$$

- When the Hue axis is polarized, it forms an angle within the orthogonalized $h - v$ plane.

HVI Transformation

Step Three：Function C_k



- Reduce ② Black Plane Noise
- We aim to automatically collapse regions of low light intensity while preserving those with higher intensity.
- Adaptive Intensity Collapse Function $\mathbf{C}_k$：

$$\mathbf{C}_k(x) = \sqrt[k]{\sin\left(\frac{\mathbf{I}_{max}(x)}{2}\right) + \varepsilon}$$

- Parameter k is trainable, ε=10^-8 is used to avoid gradient explosion.

HVI Transformation

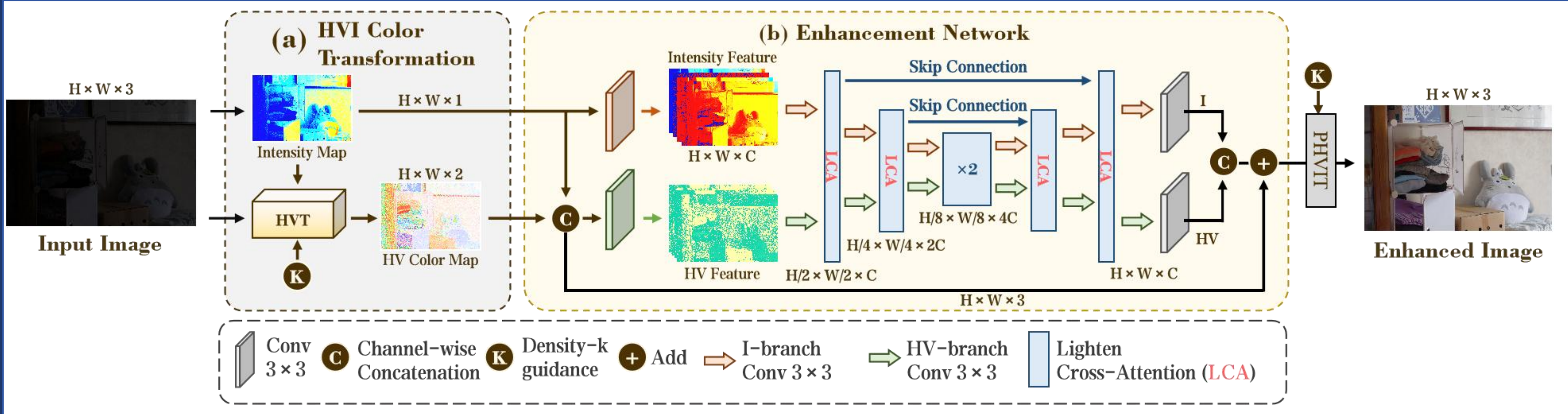
Step Four: Generate HVI Map

- 1. Generate Horizontal ($\hat{\mathbf{H}}$) map and Vertical ($\hat{\mathbf{V}}$) map as

$$\hat{\mathbf{H}} = \mathbf{C}_k \odot \mathbf{S} \odot H,$$
$$\hat{\mathbf{V}} = \mathbf{C}_k \odot \mathbf{S} \odot V,$$

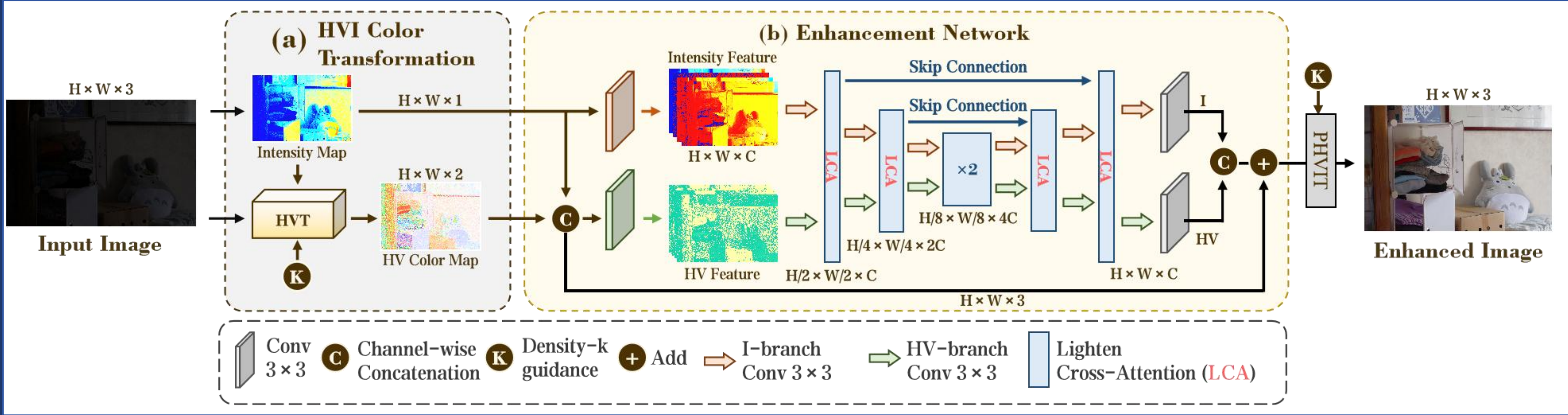
- where $h \in H$, $v \in V$ (h and v are mentioned in slide 7), and $\odot$ denotes the element-wise multiplication.
- 2. Concatenate: $\hat{\mathbf{H}}$, $\hat{\mathbf{V}}$, and $\mathbf{I}_{max}$ to form an HVI image

Color and Intensity Decoupling Network



- To more effectively utilize chromatic and brightness information in the HVI space, we introduce a novel dual-branch LLIE network, named Color and Intensity Decoupling Network (CIDNet), to separately model the HV-plane and I-axis information in the HVI space.

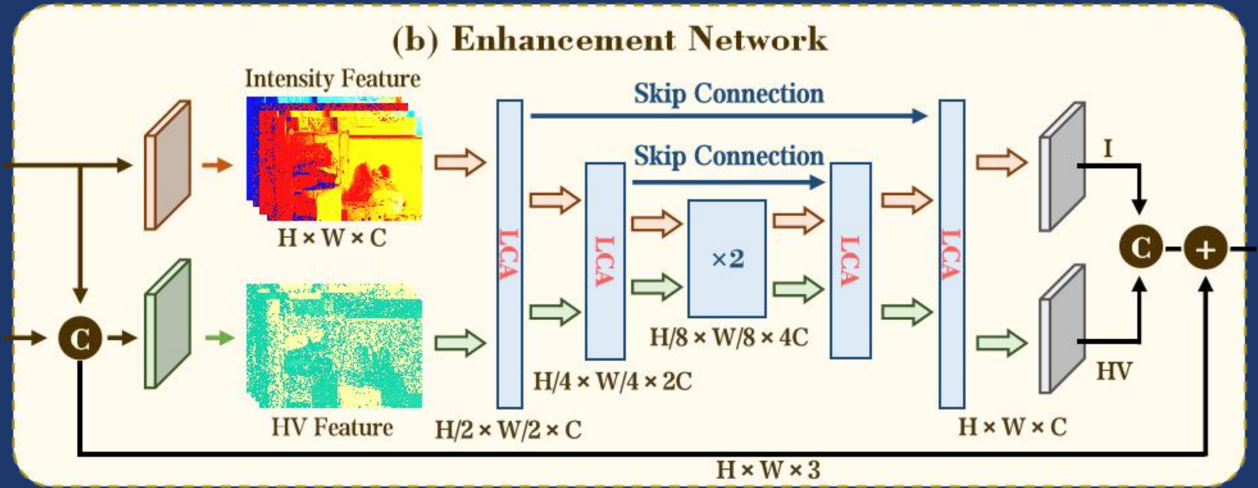
CIDNet Pipeline



- Pipeline: HVI Transformation → Enhancement Network → PHVIT
- Color Space Changes: sRGB → HVI → sRGB
- PHVIT: Perceptual-inverse HVI Transformation

Key: Dual-branch and LCA Block

- The LLIE task can be decomposed into two sub-tasks:
- (1) noise removal
- (2) brightness enhancement.



Why Dual-branch?

- These two sub-tasks follow distinct statistical patterns
- I-branch: Intensity Enhancement
- HV-branch: Remove Color Noise

Why Cross-Attention?

- **Reason 1:** The intensity is inversely proportional to image noise intensity. High-illumination requires minimal denoising and enhancement.
- **Reason 2:** Noisy- Intensity can be denoised in HV-branch.

Perceptual-inverse HVI Transformation

- Goal: HVI Map → sRGB Image
- Step One: HVI → HSV

$$\hat{h} = \frac{\hat{\mathbf{H}}}{\mathbf{C}_k + \varepsilon}, \hat{v} = \frac{\hat{\mathbf{V}}}{\mathbf{C}_k + \varepsilon},$$

Mentioned at Slide 8 and Slide 9

$$\mathbf{H} = \arctan\left(\frac{\hat{v}}{\hat{h}}\right) \mod 1,$$

$$\mathbf{S} = \alpha_S \sqrt{\hat{h}^2 + \hat{v}^2},$$

$$\mathbf{V} = \alpha_I \hat{\mathbf{I}}_{\mathbf{I}},$$

α_S and α_I are the customizing linear parameters to change the image color saturation and brightness.

- Step Two: HSV → sRGB

James D Foley and Andries Van Dam. *Fundamentals of interactive computer graphics.* Addison-Wesley Longman Publishing Co., Inc., 1982.

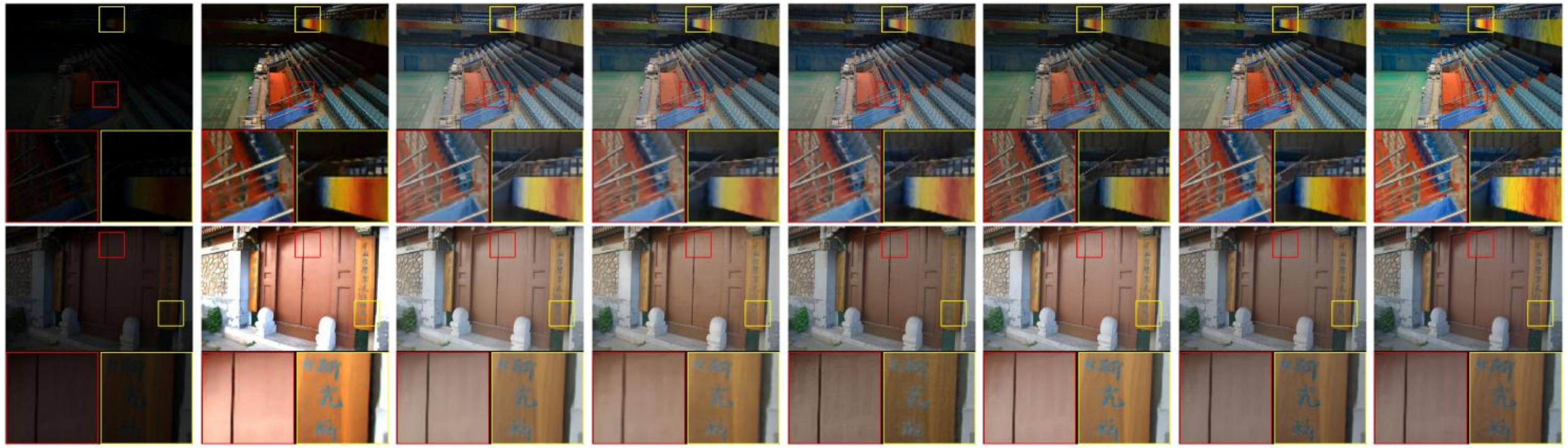
Loss Function

Loss in HVI Color Space, λ is a eighting hyperparameter to balance the losses in the two different color spaces.

$$L = \lambda \cdot l(\hat{\mathbf{I}}_{\mathbf{HVI}}, \mathbf{I}_{\mathbf{HVI}}) + l(\hat{\mathbf{I}}, \mathbf{I}),$$

Loss in sRGB Color Space.

Results on LOL Datasets



Input RUAS [33] LLFlow [41] SNRNet [46] RetFormer [3] GSAD [15] CIDNet(Ours) GroundTruth

Methods	Color Model	Complexity		LOLv1			LOLv2-Real			LOLv2-Synthetic		
		Params/M	FLOPs/G	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓
RetinexNet [43]	Retinex	0.84	584.47	18.915	0.427	0.470	16.097	0.401	0.543	17.137	0.762	0.255
KinD [51]	Retinex	8.02	34.99	23.018	0.843	0.156	17.544	0.669	0.375	18.320	0.796	0.252
ZeroDCE [10]	RGB	0.075	4.83	21.880	0.640	0.335	16.059	0.580	0.313	17.712	0.815	0.169
RUAS [33]	Retinex	0.003	0.83	18.654	0.518	0.270	15.326	0.488	0.310	13.765	0.638	0.305
LLFlow [41]	RGB	17.42	358.4	24.998	0.871	0.117	17.433	0.831	0.176	24.807	0.919	0.067
EnlightenGAN [18]	RGB	114.35	61.01	20.003	0.691	0.317	18.230	0.617	0.309	16.570	0.734	0.220
SNR-Aware [46]	SNR+RGB	4.01	26.35	26.716	0.851	0.152	21.480	0.849	0.163	24.140	0.928	0.056
Bread [11]	YCbCr	2.02	19.85	25.299	0.847	0.155	20.830	0.847	0.174	17.630	0.919	0.091
PairLIE [8]	Retinex	0.33	20.81	23.526	0.755	0.248	19.885	0.778	0.317	19.074	0.794	0.230
LLFormer [39]	RGB	24.55	22.52	25.758	0.823	0.167	20.056	0.792	0.211	24.038	0.909	0.066
RetinexFormer [3]	Retinex	1.53	15.85	27.140	0.850	0.129	22.794	0.840	0.171	25.670	0.930	0.059
GSAD [16]	RGB	17.36	442.02	27.605	0.876	0.092	20.153	0.846	0.113	24.472	0.929	0.051
QuadPrior [40]	Kubelka-Munk	1252.75	1103.20	22.849	0.800	0.201	20.592	0.811	0.202	16.108	0.758	0.114
CIDNet(Ours)	HVI	1.88	7.57	28.201	0.889	0.079	24.111	0.871	0.108	25.705	0.942	0.045

Results on SICE and Sony-total-Dark

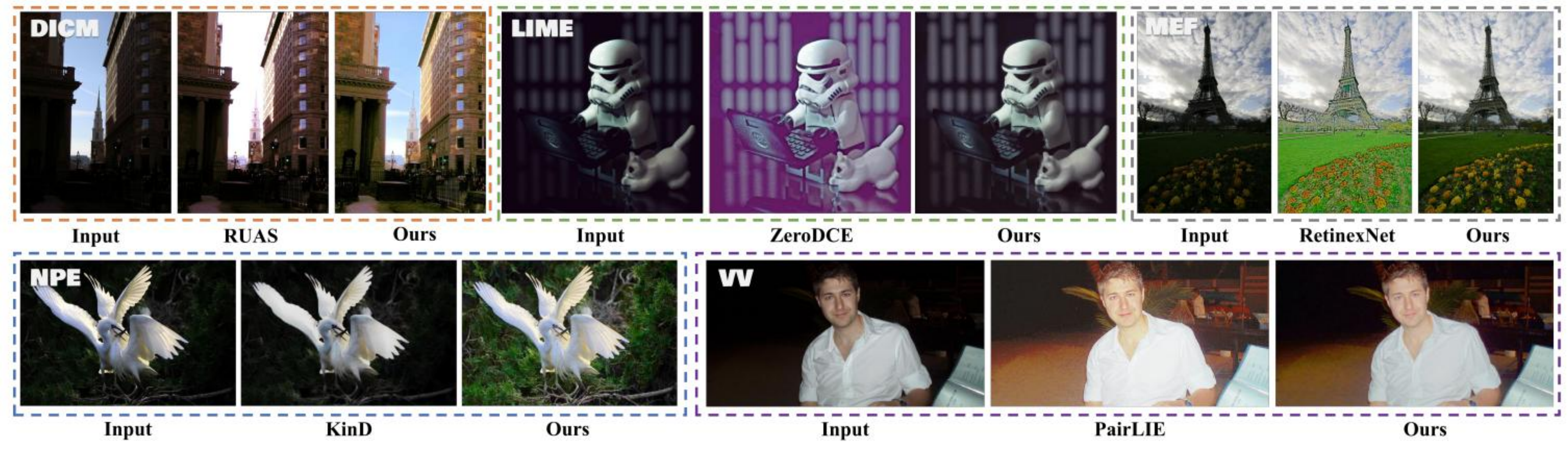
Methods	SICE		Sony-Total-Dark	
	PSNR↑	SSIM↑	PSNR↑	SSIM↑
RetinexNet [43]	12.424	0.613	15.695	0.395
ZeroDCE [10]	12.452	0.639	14.087	0.090
URetinexNet [44]	10.899	0.605	15.519	0.323
RUAS [33]	8.656	0.494	12.622	0.081
LLFlow [41]	12.737	0.617	16.226	0.367
CIDNet (Ours)	13.435	0.642	22.904	0.676

Visualization on Sony-Total-Dark

Results on Five Unpaired Dataset

Methods	Unpaired	
	BRIS↓	NIQE↓
RetinexNet [43]	23.286	4.558
ZeroDCE [10]	26.343	4.763
URetinexNet [44]	26.359	3.829
RUAS [33]	36.372	4.800
LLFlow [41]	28.087	4.221
CIDNet (Ours)	23.521	3.523

While CIDNet does not outperform RetinexNet in the BRISQUE metric, its recovered perceptual results are closer to realistic appearances than RetinexNet.



Generalizing HVI to Other LLIE Models

Methods	FourLLIE [37]	LEDNet [57]	SNR-Aware [46]	LLFormer [39]	GSAD [15]	DiffLight [6]	**CIDNet**
PSNR↑	22.730(+0.381)	23.394(+3.456)	22.251(+0.771)	22.671(+2.615)	23.715(+3.562)	23.969(+1.364)	**24.111**
SSIM↑	0.856(+0.009)	0.837(+0.010)	0.840(-0.009)	0.852(+0.060)	**0.876**(+0.030)	0.859(+0.003)	0.871
LPIPS↓	0.125(+0.011)	0.115(-0.005)	0.117(-0.054)	0.117(-0.094)	**0.103**(-0.010)	0.109(-0.012)	0.108
GPU Time/s↓	0.075	0.054	0.070	0.139	0.315	0.578	**0.053**
Model Type	CNN	CNN	Transformer	Transformer	Diffusion	CNN+Diffusion	Transformer

- To verify the effectiveness of the HVI color space, we further evaluate its performance when it is used with different sRGB-based LLIE models.
- Integrate HVIT and PHVIT as a plug-and-play module into six SOTA methods.
- Transforming to the HVI color space improves PSNR, SSIM, and LPIPS metrics across various methods compared to the results in the sRGB color space.

Ablation 1: Color Spaces

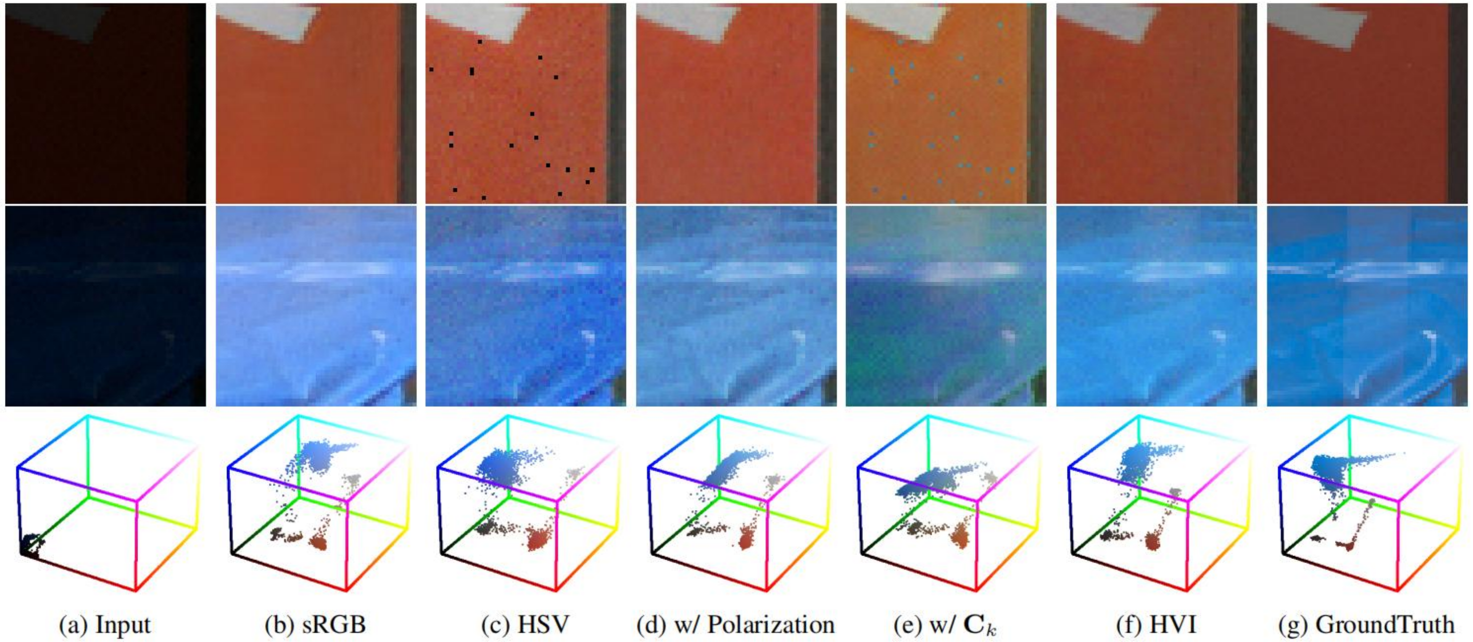
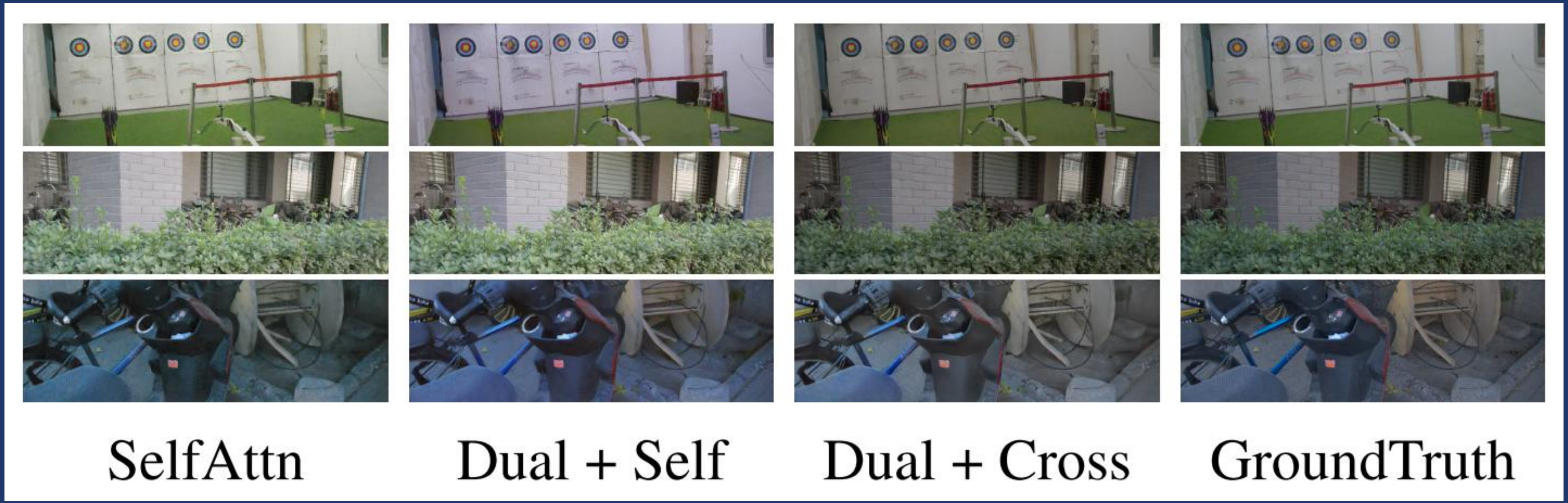


Figure 5. Top and middle rows are ablation results on LOLv2-Real for five different color spaces used by CIDNet. The bottom row provides a visual comparison by mapping the pixel values of the results to sRGB. Note that due to the dual-branch and the cross-attention mechanism are specifically designed for HVI, we only use UNet [31] with self attentions [49] for a fair comparison.

Metrics		PSNR↑	SSIM↑	LPIPS↓
Color Space	sRGB	20.062	0.825	0.137
	HSV	21.349	0.801	0.167
	HVI (w/ Polarization Only)	21.558	0.821	0.149
	HVI (w/ $\mathbf{C}_k$ Only)	21.536	0.825	0.179
Full Model (HVI-CIDNet)		24.111	0.871	0.108

Ablation 2: Different Module



Metrics		PSNR↑	SSIM↑	LPIPS↓
Structure	UNet Baseline [31]	19.306	0.778	0.222
	SelfAttn [49]	22.313	0.835	0.126
	Dual+SelfAttn [49]	23.159	0.856	0.116
Full Model (HVI-CIDNet)		24.111	0.871	0.108

Dual-branch and Cross Attention are all beneficial to the task.

Ablation 3: Loss Function

Metrics		PSNR↑	SSIM↑	LPIPS↓
Loss	HVI Only	23.221	0.854	0.132
	sRGB Only	23.319	0.857	0.123
Full Model (HVI-CIDNet)		24.111	0.871	0.108

- The HVI loss lacks pixel-level spatial consistency constraints, leading to a loss of structural detail in the image and thus lower performance across the three metrics, especially in the LPIPS metric.
- Using only sRGB loss is focused on pixel-space enhancement, neglecting the low-light probability distribution in the HVI color space, resulting in undesired color imbalance.
- Conclusion: HVI + sRGB performs the best score.

Thanks For Listening

Contact: Yixu Feng (yixu-nwpu@mail.nwpu.edu.cn)