Supplementary Materials for Learning Structural Illumination for Unsupervised Low-Light Enhancement

A.1 Relative Illumination beyond Brightness Ordering

We clarify why RISE models continuous, spatially conditioned illumination relations instead of directly adopting the regional brightness ordering observed in a low-light input. The analysis first examines whether the input ordering reliably describes the desired normal-light appearance, and then highlights the additional information captured by our relative illumination formulation.

The relative brightness ordering in a low-light observation is not necessarily consistent with that in its normally exposed counterpart. Regional luminance is not a direct measurement of illumination, but is jointly determined by illumination, surface reflectance, sensor noise, quantization, and the camera response. In particular, weak measurements in severely underexposed regions can compress or even reverse pairwise brightness relations. Directly treating the input ordering as the target illumination structure may therefore impose incorrect ordinal constraints on the enhanced result. We analyze this phenomenon using paired low-light and ground-truth images only for evaluation; the ground truth is never used to train RISE.

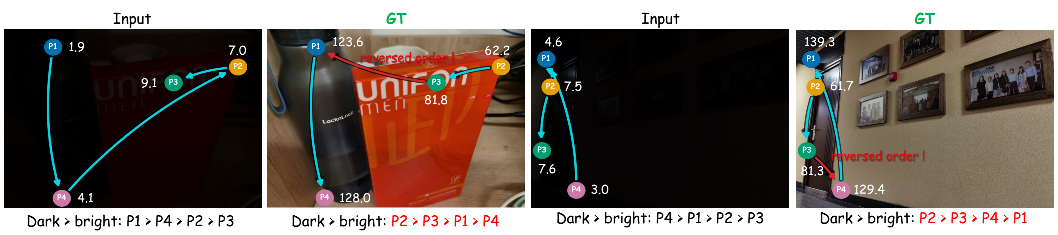

Figure 1: Inconsistency of relative brightness ordering. Representative examples where the regional brightness ordering in low-light inputs differs from that in the corresponding ground-truth images, indicating that input brightness order is not a reliable target illumination constraint.

Table 1: Quantitative comparison of luminance distributions. Values represent the Wasserstein Distance (↓) to the corresponding ground truth. The best and second-best results are denoted by boldface and underlining, respectively.

Method	LOL-v1	LOLv2-Real	LOLv2-Syn	LSRW-Huawei	LSRW-Nikon	UHD-LL
Zero-DCE (Guo et al., 2020)	0.1684	0.1070	0.0907	0.1223	0.0923	0.0744
RUAS (Liu et al., 2021)	0.0792	0.0890	0.1022	0.0928	0.1300	0.2075
SCI (Ma et al., 2022)	0.1816	0.1179	0.0618	0.1402	0.0572	0.1071
CLIP-LIT (Liang et al., 2023)	0.2374	0.1587	0.1325	0.1797	0.1589	0.1520
CoLIE (Chobola et al., 2024)	0.2087	0.1643	0.1961	0.1623	0.1922	0.1414
CLODE (Jung et al., 2025)	**0.0266**	0.0503	0.0636	0.0475	0.0829	0.1186
RISE	0.0769	**0.0213**	0.0348	**0.0305**	**0.0281**	**0.0271**
RISE†	0.0609	0.0422	**0.0345**	0.0351	**0.0281**	0.0525

These inconsistencies indicate that input brightness relations should serve as evidence for illumination estimation rather than constraints to be directly preserved. A rank-based representation retains only whether one region is brighter than another. In contrast, the discrepancy used by RISE satisfies $\delta_n(p) = Z_{k_n} - Z_p = \log(P_{k_n}/P_p)$, preserving the magnitude of the relative brightness ratio while remaining invariant to global intensity scaling. RISE further combines this continuous relation with spatial attenuation $\rho_n(p)$, multiple KICs, predicted KIC luminance, and scene context to construct a scene-conditioned gain field. Importantly, KICs are selected as high-SNR anchors rather than fixed ordering targets. Their relations guide the estimated correction without enforcing the input brightness ordering on the enhanced output. Consequently, RISE can distinguish illumination patterns that share the same ordinal ranking but exhibit different attenuation strengths, while correcting unreliable relations inherited from low-SNR observations.

A.2 DETAILS OF THE TRANSITION FROM LOG-GAIN TO IMAGE-SPACE ENHANCEMENT

We further clarify the transition from Eq. (4) to Eq. (5) in the main paper. RISE does not predict an additive offset to the image intensity. Instead, it predicts the logarithm of a multiplicative exposure gain. At the patch level, Eq. (4) decomposes this log-gain into relative illumination structure and absolute exposure:

$$\hat{G}(p) = S(p) + \eta.$$

Before it is applied to the input image, the patch-level field is upsampled to the pixel grid and mapped back to the linear-gain domain through the exponential function. For a pixel x, the resulting gain is

$$\begin{aligned} g(x) &= \exp\Big(\mathrm{Up}(\hat{\mathbf{G}})(x)\Big) \\ &= \exp(\mathrm{Up}(\mathbf{S})(x) + \eta) \\ &= \exp(\mathrm{Up}(\mathbf{S})(x)) \exp(\eta). \end{aligned}$$

The enhanced image in Eq. (5) is therefore obtained by multiplying this gain with the low-light input:

$$\hat{\mathbf{I}}(x) = \mathbf{I}_{\mathrm{low}}(x) \odot g(x).$$

This derivation explains why the two components are additive in log-gain space but multiplicative in the linear-gain domain: the relative term determines the spatially varying correction, whereas η scales the exposure of the entire image. The parameterization also gives a direct interpretation of the predicted values. Specifically, $\hat{G} = 0$ yields $g = 1$ and leaves the intensity unchanged; $\hat{G} = \log 2$ yields $g = 2$ and doubles it; and $\hat{G} = \log 0.5$ yields $g = 0.5$ and halves it. More generally, positive log-gain values enhance exposure, while negative values suppress it.

A.3 QUANTITATIVE LUMINANCE DISTRIBUTION COMPARISON

Beyond standard image restoration metrics, we further examine the luminance distributions produced by different unsupervised LLIE methods. We compute the Wasserstein Distance between the luminance distribution of each enhanced result and that of the corresponding ground truth. A smaller distance indicates that the enhanced image better matches the target exposure distribution. As shown in Table 1, RISE obtains the lowest or near-lowest distance on most benchmarks, suggesting that the proposed illumination modeling yields more faithful luminance distributions across diverse datasets.

A.4 MORE QUALITATIVE RESULTS

Additional qualitative results on LOL, LSRW, UHD-LL, and SICE complement the main comparisons. These examples cover diverse illumination distributions and emphasize both the recovery of poorly illuminated content and the preservation of details in bright regions.

A.5 MORE REAL-WORLD GENERALIZATION RESULTS

For out-of-benchmark generalization, we apply the same LSRW-Nikon checkpoint to low-light photographs captured with an iPhone 16 Pro, without additional training or scene-specific tuning.

A.6 FAILURE CASES

RISE estimates illumination from the visual evidence retained in the input and cannot reconstruct scene content that falls below the sensor noise floor. Representative examples illustrate this limitation in extremely dark regions.

Figure 2: Additional visual comparisons on the LOL benchmarks.

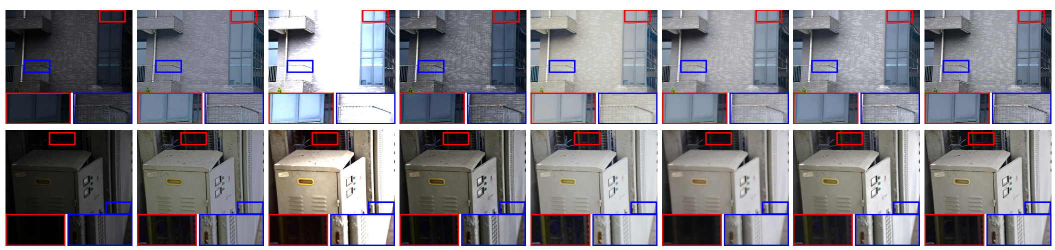

Figure 3: Additional visual comparisons on LSRW.

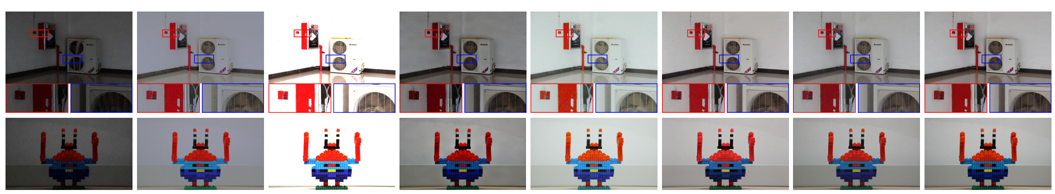

Figure 4: Additional visual comparisons on UHD-LL.

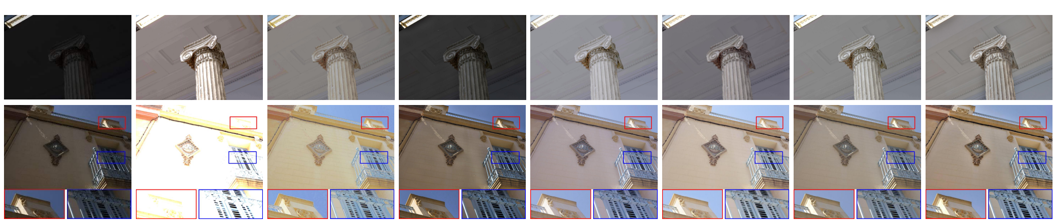

Figure 5: Additional visual comparisons on SICE.

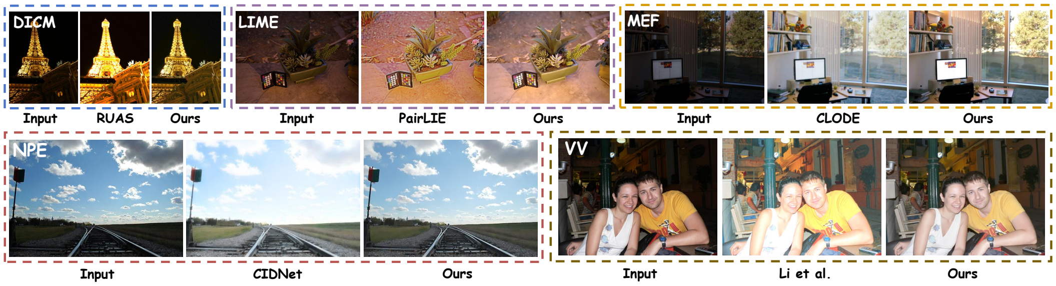

Figure 6: Visual comparisons on No-Reference Datasets.

Figure 7: Additional comparisons on real-world low-light scenes. RISE provides balanced exposure in both outdoor and indoor scenes while avoiding severe highlight saturation. All results are produced by the LSRW-Nikon checkpoint without additional tuning.

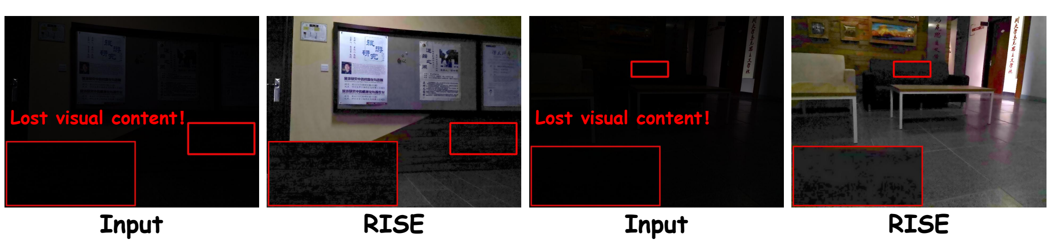

Figure 8: Failure cases under unrecoverable signal loss. RISE improves overall visibility but cannot recover structures absent from the recorded input and may retain residual noise in these regions. Red boxes and enlarged crops indicate the affected areas.

REFERENCES

Tomáš Chobola, Yu Liu, Hanyi Zhang, Julia A Schnabel, and Tingying Peng. Fast context-based low-light image enhancement via neural implicit representations. In *European Conference on Computer Vision*, pp. 413–430. Springer, 2024.

Chunle Guo, Chongyi Li, Jichang Guo, Chen Change Loy, Junhui Hou, Sam Kwong, and Runmin Cong. Zero-reference deep curve estimation for low-light image enhancement. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 1780–1789, 2020.

Donggoo Jung, Daehyun Kim, and Tae Hyun Kim. Continuous exposure learning for low-light image enhancement using neural odes. In *The Thirteenth International Conference on Learning Representations*, 2025.

Zhexin Liang, Chongyi Li, Shangchen Zhou, Ruicheng Feng, and Chen Change Loy. Iterative prompt learning for unsupervised backlit image enhancement. In *Proceedings of the IEEE/CVF international conference on computer vision*, pp. 8094–8103, 2023.

Risheng Liu, Long Ma, Jiaao Zhang, Xin Fan, and Zhongxuan Luo. Retinex-inspired unrolling with cooperative prior architecture search for low-light image enhancement. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 10561–10570, 2021.

Long Ma, Tengyu Ma, Risheng Liu, Xin Fan, and Zhongxuan Luo. Toward fast, flexible, and robust low-light image enhancement. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 5637–5646, 2022.