

PercepLIE: A New Path to Perceptual Low-Light Image Enhancement —Supplementary Materials—

Cong Wang

Shenzhen Campus of Sun Yat-sen University & The Hong
Kong Polytechnic University
Shenzhen & Hong Kong, China
supercong94@gmail.com

Jie Mu

Dongbei University of Finance and Economics
Dalian, China
jiemu@dufe.edu.cn

Chengjin Yu

Anhui University
Hefei, China
23073@ahu.edu.cn

Wei Wang*

Shenzhen Campus of Sun Yat-sen University
Shenzhen, China
wangwei29@mail.sysu.edu.cn

ABSTRACT

In this document, we provide the detailed architecture of our network in Sec. 1. In Sec 2, we provide more visual comparisons of our PercepLIE and state-of-the-art (SOTA) ones. Sec. 3 presents some examples of estimated multiple luminance details and their weight maps in the Multi-luminance Detail Compensation (MDC) and the exploitable adjustment image and the corresponding weight map in the Global Color Adjustment (GCA).

CCS CONCEPTS

• Computing methodologies → Computational photography.

KEYWORDS

Low-Light Image Enhancement, Latent Light Representation, Multi-luminance Detail Compensation, Global Color Adjustment

ACM Reference Format:

Cong Wang, Chengjin Yu, Jie Mu, and Wei Wang. 2024. PercepLIE: A New Path to Perceptual Low-Light Image Enhancement —Supplementary Materials—. In *Proceedings of the 32nd ACM International Conference on Multimedia (MM '24)*, October 28–November 1, 2024, Melbourne, VIC, Australia. ACM, Ontario, Canada, 13 pages. <https://doi.org/10.1145/3664647.3681399>

1 NETWORK ARCHITECTURE

The architectures of the encoder and decoder in VQGAN at **Stage 1** are detailed in Tabs. 1, 2, respectively. Tab. 3 details the architecture of the codebook. Following [1], the codebook has 1024 codes and each code is a vector with 512 dimensions. The architectures of the MDC decoder (MDC-Dec) and GCA decoder (GCA-Dec) are

detailed in Tabs. 4, 5, respectively. The triple feature transformation module (TriFTM) is only used in **Stage 3** for fusing the features in the decoder of VQGAN, MDC-Dec, and GCA-Dec, which is not shown in Tab. 2.

2 MORE COMPARISON RESULTS ON REAL-WORLD DATASETS

Figures 1-8 show more real-world comparison results. These additional examples demonstrate that our PercepLIE is able to generate more natural and clearer results with finer structures.

3 ESTIMATED RESULTS IN MDC AND GCA

In the paper, we utilize the MDC and GCA to estimate multiple luminance details and the exploitable adjustment image and corresponding weight maps. In Figures 9-10, we show some estimated results in MDC and GCA. These results demonstrate that our MDC is able to generate better multiple luminance details and the GCA is able to produce better exploitable adjustment images, thus leading to our better enhancement results.

*Wei Wang is the Corresponding author.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
MM '24, October 28–November 1, 2024, Melbourne, VIC, Australia.

© 2024 Copyright held by the owner/author(s). Publication rights licensed to ACM.
ACM ISBN 979-8-4007-0686-8/24/10...\$15.00
<https://doi.org/10.1145/3664647.3681399>

Table 1: Encoder.

Layer	Encoder1	Downsample	Encoder2	Downsample	Encoder3	Downsample	Encoder4	Encoder5
Shape	$H \times W \times 64$	$\frac{H}{2} \times \frac{W}{2} \times 128$	$\frac{H}{2} \times \frac{W}{2} \times 128$	$\frac{H}{4} \times \frac{W}{4} \times 256$	$\frac{H}{4} \times \frac{W}{4} \times 256$	$\frac{H}{8} \times \frac{W}{8} \times 256$	$\frac{H}{8} \times \frac{W}{8} \times 256$	$\frac{H}{8} \times \frac{W}{8} \times 512$
Blocks	3 ResBlocks	-	2 ResBlocks	-	2 ResBlocks	-	2 ResBlocks	1 Conv
Kernel sizes	3×3	-	3×3	-	3×3	-	3×3	3×3
Stride	1	-	1	-	1	-	1	1

Table 2: Decoder.

Layer	Decoder1	Decoder2	Upsample	Decoder3	Upsample	Decoder4	Upsample	Decoder5	Decoder6
Shape	$\frac{H}{8} \times \frac{W}{8} \times 512$	$\frac{H}{8} \times \frac{W}{8} \times 256$	$\frac{H}{4} \times \frac{W}{4} \times 256$	$\frac{H}{4} \times \frac{W}{4} \times 256$	$\frac{H}{2} \times \frac{W}{2} \times 256$	$\frac{H}{2} \times \frac{W}{2} \times 128$	$H \times W \times 128$	$H \times W \times 64$	$H \times W \times 3$
Blocks	1 Conv	2 ResBlocks	-	2 ResBlocks	-	2 ResBlocks	-	3 ResBlocks	1 Conv
Kernel sizes	3×3	3×3	-	3×3	-	3×3	-	3×3	3×3
Stride	1	1	-	1	-	1	-	1	1

Table 3: Codebook. The codebook has 1024 discrete codes and each code is a vector with 512 dimensions.

Layer	Codebook
Shape	1024×512
Blocks	nn.Embedding (in PyTorch)

Table 4: Decoder of MDC. The MDC outputs estimated three luminance details $S \in \mathbb{R}^{H \times W \times 3}$ and corresponding weight maps $M \in \mathbb{R}^{H \times W \times 1}$.

Layer	Decoder1	Decoder2	Upsample	Decoder3	Upsample	Decoder4	Upsample	Decoder5	Decoder6
Shape	$\frac{H}{8} \times \frac{W}{8} \times 512$	$\frac{H}{8} \times \frac{W}{8} \times 256$	$\frac{H}{4} \times \frac{W}{4} \times 256$	$\frac{H}{4} \times \frac{W}{4} \times 256$	$\frac{H}{2} \times \frac{W}{2} \times 256$	$\frac{H}{2} \times \frac{W}{2} \times 128$	$H \times W \times 128$	$H \times W \times 64$	$H \times W \times 3$ $H \times W \times 1$ $H \times W \times 3$ $H \times W \times 1$ $H \times W \times 3$ $H \times W \times 1$
Blocks	1 Conv	2 ResBlocks	-	2 ResBlocks	-	2 ResBlocks	-	3 ResBlocks	1 Conv 1 Conv 1 Conv 1 Conv 1 Conv 1 Conv
Kernel sizes	3×3	3×3	-	3×3	-	3×3	-	3×3	3×3
Stride	1	1	-	1	-	1	-	1	1

Table 5: Decoder of GCA. The GCA outputs exploitable adjustment image $Y_{GCA} \in \mathbb{R}^{H \times W \times 3}$ and corresponding weight map $M_{GCA} \in \mathbb{R}^{H \times W \times 1}$.

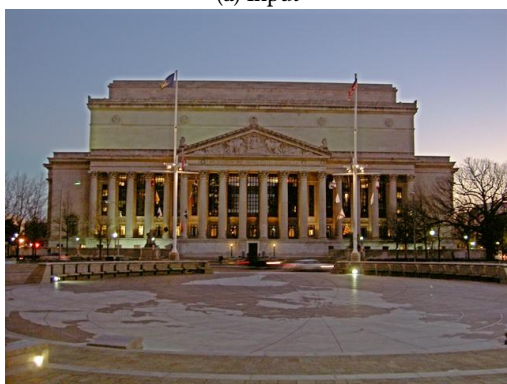
Layer	Decoder1	Decoder2	Upsample	Decoder3	Upsample	Decoder4	Upsample	Decoder5	Decoder6
Shape	$\frac{H}{8} \times \frac{W}{8} \times 512$	$\frac{H}{8} \times \frac{W}{8} \times 256$	$\frac{H}{4} \times \frac{W}{4} \times 256$	$\frac{H}{4} \times \frac{W}{4} \times 256$	$\frac{H}{2} \times \frac{W}{2} \times 256$	$\frac{H}{2} \times \frac{W}{2} \times 128$	$H \times W \times 128$	$H \times W \times 64$	$H \times W \times 3$ $H \times W \times 1$
Blocks	1 Conv	2 ResBlocks	-	2 ResBlocks	-	2 ResBlocks	-	3 ResBlocks	1 Conv 1 Conv
Kernel sizes	3×3	3×3	-	3×3	-	3×3	-	3×3	3×3
Stride	1	1	-	1	-	1	-	1	1



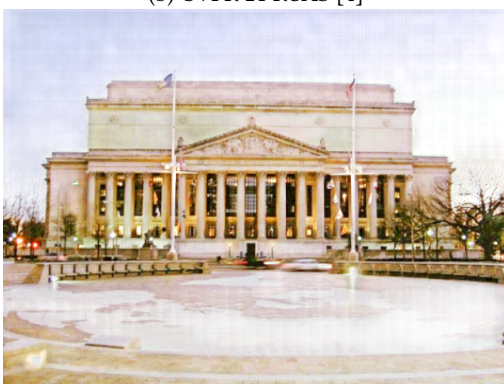
(a) Input



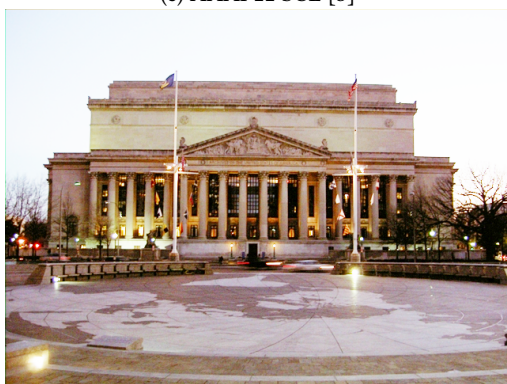
(b) CVPR'21 RUAS [4]



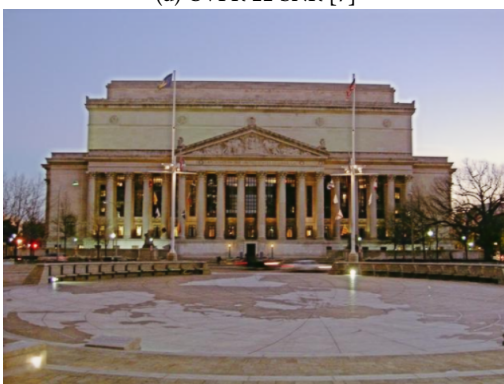
(c) AAAI'22 SCL [3]



(d) CVPR'22 SNR [7]



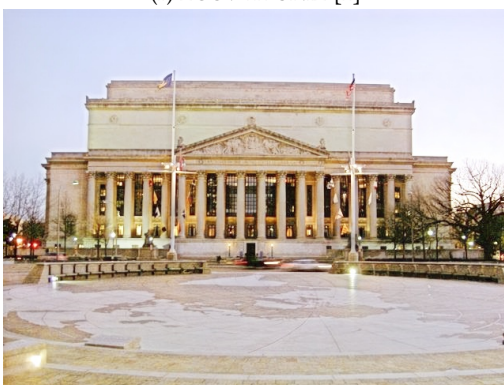
(e) CVPR'22 SCI [5]



(f) ECCV'22 UNIE [2]



(g) AAAI'22 LLFlow [6]



(h) PercepLIE

Figure 1: Visual comparisons with recent SOTAs on real-world datasets. Our PercepLIE is capable of producing a clearer and cleaner result.

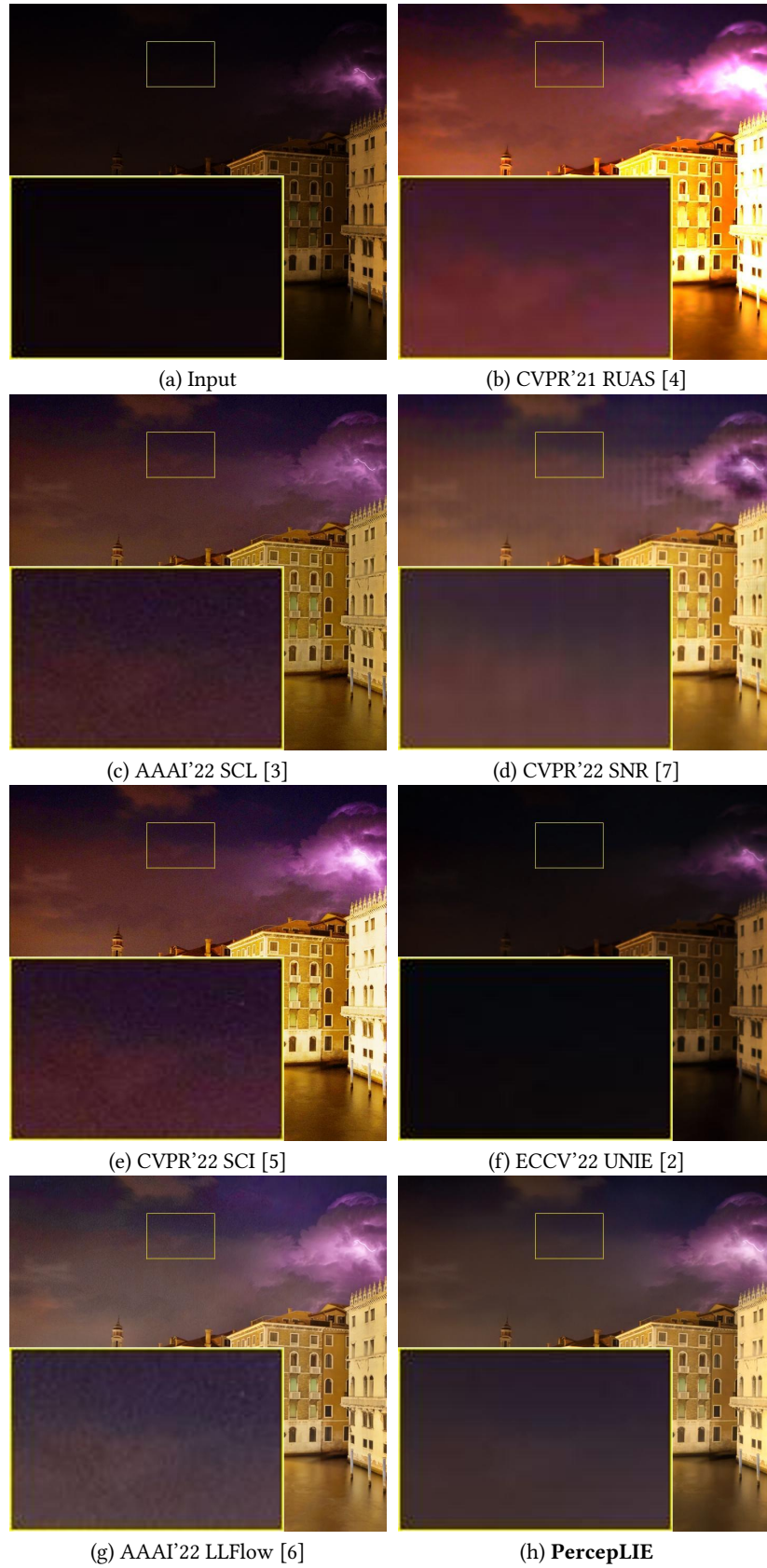


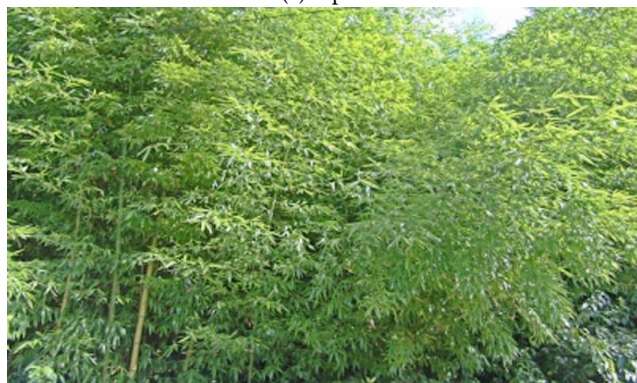
Figure 2: Visual comparisons with recent SOTAs on real-world datasets. Our PercepLIE is able to produce a clearer result.



(a) Input



(b) CVPR'21 RUAS [4]



(c) AAAI'22 SCL [3]



(d) CVPR'22 SNR [7]



(e) CVPR'22 SCI [5]



(f) ECCV'22 UNIE [2]

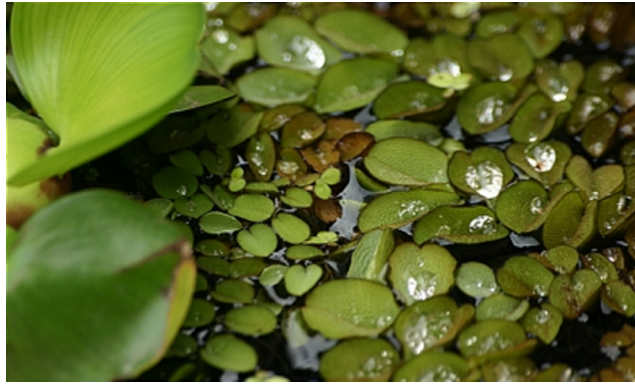


(g) AAAI'22 LLFlow [6]



(h) PercepLIE

Figure 3: Visual comparisons with recent SOTAs on real-world datasets. Our PercepLIE is able to generate a more natural result with better colors and details.



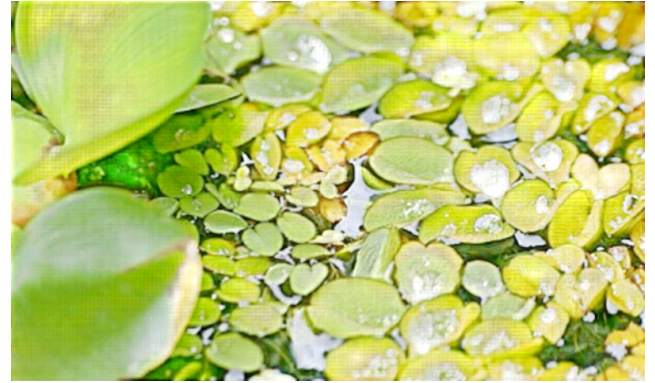
(a) Input



(b) CVPR'21 RUAS [4]



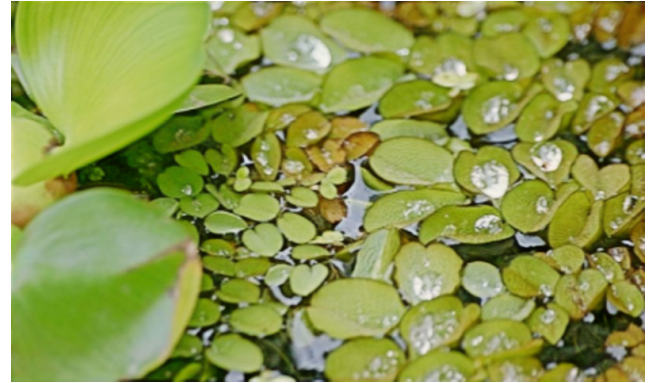
(c) AAAI'22 SCL [3]



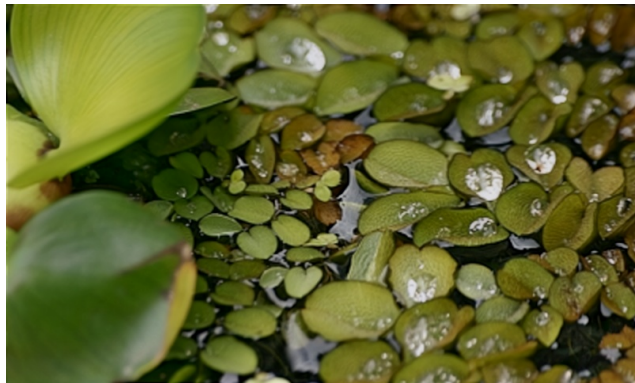
(d) CVPR'22 SNR [7]



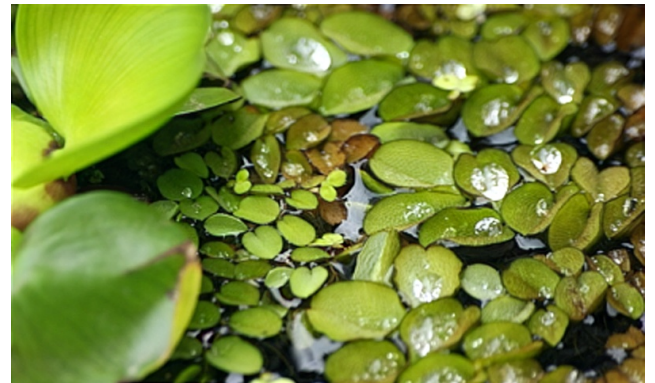
(e) CVPR'22 SCI [5]



(f) ECCV'22 UNIE [2]



(g) AAAI'22 LLFlow [6]



(h) PercepLIE

Figure 4: Visual comparisons with recent SOTAs on real-world datasets. Our PercepLIE is able to generate a more natural result with better colors and details.



(a) Input



(b) CVPR'21 RUAS [4]



(c) AAAI'22 SCL [3]



(d) CVPR'22 SNR [7]



(e) CVPR'22 SCI [5]



(f) ECCV'22 UNIE [2]



(g) AAAI'22 LLFlow [6]



(h) PercepLIE

Figure 5: Visual comparisons with recent SOTAs on real-world datasets. Our PercepLIE is able to generate a more natural result with better colors and details.



(a) Input



(b) CVPR'21 RUAS [4]



(c) AAAI'22 SCL [3]



(d) CVPR'22 SNR [7]



(e) CVPR'22 SCI [5]



(f) ECCV'22 UNIE [2]



(g) AAAI'22 LLFlow [6]



(h) PercepLIE

Figure 6: Visual comparisons with recent SOTAs on real-world datasets. Our PercepLIE is able to generate a more natural result with better colors and details.



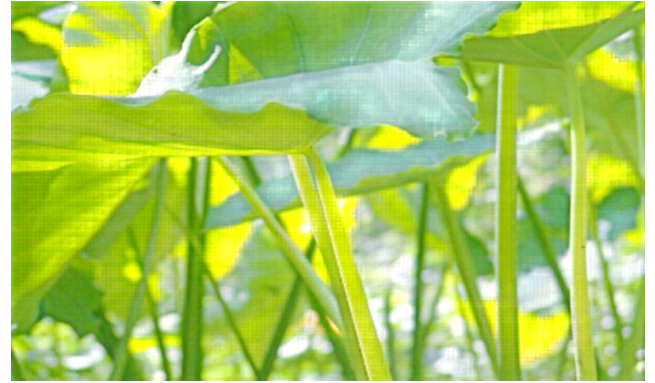
(a) Input



(b) CVPR'21 RUAS [4]



(c) AAAI'22 SCL [3]



(d) CVPR'22 SNR [7]



(e) CVPR'22 SCI [5]



(f) ECCV'22 UNIE [2]



(g) AAAI'22 LLFlow [6]



(h) PercepLIE

Figure 7: Visual comparisons with recent SOTAs on real-world datasets. Our PercepLIE is able to generate a more natural result with better colors and details.



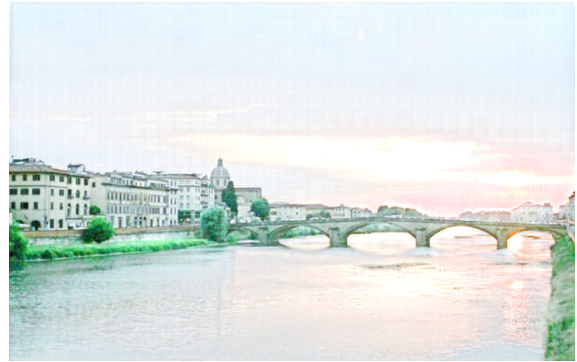
(a) Input



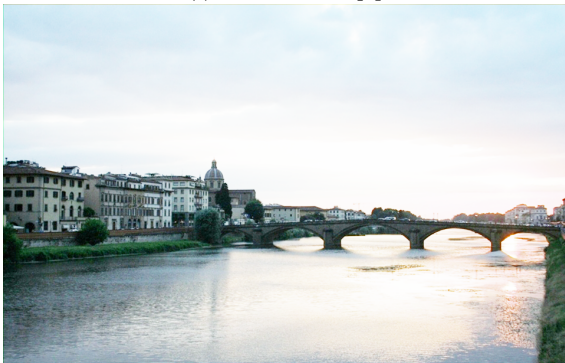
(b) CVPR'21 RUAS [4]



(c) AAAI'22 SCL [3]



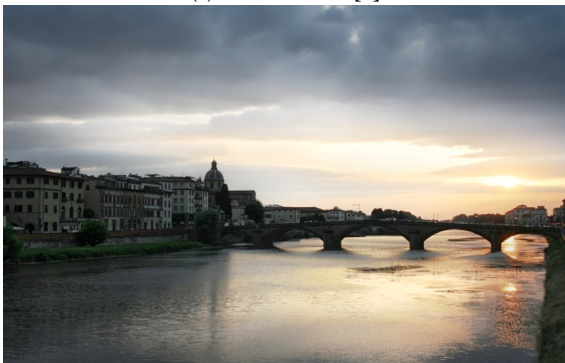
(d) CVPR'22 SNR [7]



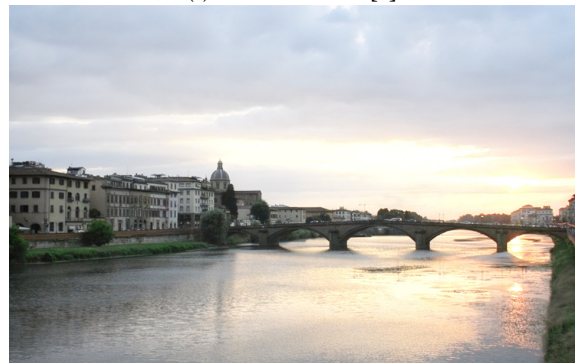
(e) CVPR'22 SCI [5]



(f) ECCV'22 UNIE [2]



(g) AAAI'22 LLFlow [6]



(h) PercepLIE

Figure 8: Visual comparisons with recent SOTAs on real-world datasets. The α is set as 1. Our PercepLIE is able to generate a more natural result with better colors and details.

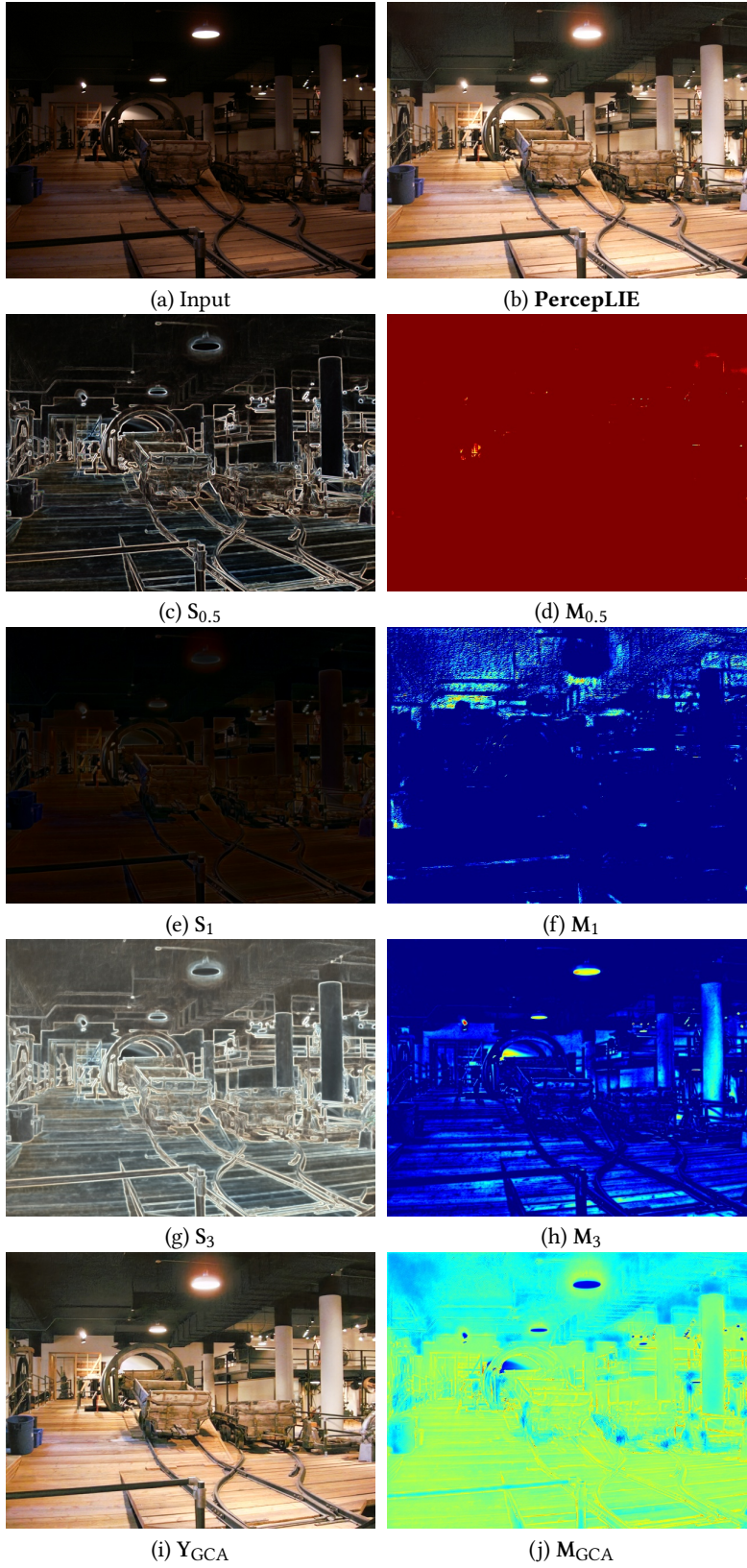


Figure 9: $S \in \mathbb{R}^{H \times W \times 3}$ denotes the estimated multiple luminance details and corresponding weight maps $M \in \mathbb{R}^{H \times W \times 1}$ in MDC. The $Y_{GCA} \in \mathbb{R}^{H \times W \times 3}$ and $M_{GCA} \in \mathbb{R}^{H \times W \times 1}$ denote the exploitable adjustment image and corresponding weight map in GCA.

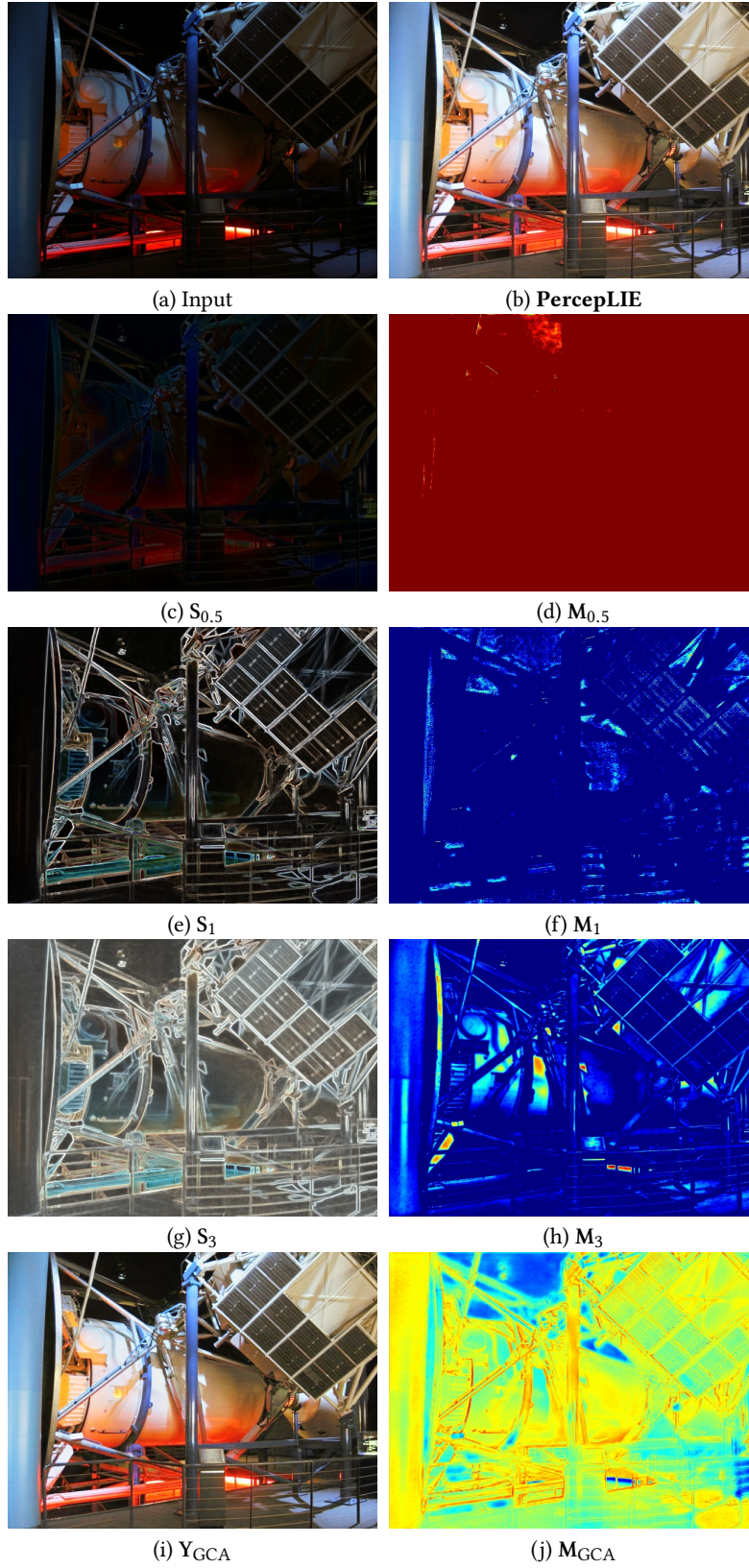


Figure 10: $S \in \mathbb{R}^{H \times W \times 3}$ denotes the estimated multiple luminance details and corresponding weight maps $M \in \mathbb{R}^{H \times W \times 1}$ in MDC. The $Y_{GCA} \in \mathbb{R}^{H \times W \times 3}$ and $M_{GCA} \in \mathbb{R}^{H \times W \times 1}$ denote the exploitable adjustment image and corresponding weight map in GCA.

REFERENCES

- [1] Chaofeng Chen, Xinyu Shi, Yipeng Qin, Xiaoming Li, Xiaoguang Han, Tao Yang, and Shihui Guo. Real-world blind super-resolution via feature matching with implicit high-resolution priors. In *ACM MM*, pages 1329–1338, 2022.
- [2] Yeying Jin, Wenhan Yang, and Robby T. Tan. Unsupervised night image enhancement: When layer decomposition meets light-effects suppression. In *ECCV*, volume 13697, pages 404–421, 2022.
- [3] Dong Liang, Ling Li, Mingqiang Wei, Shuo Yang, Liyan Zhang, Wenhan Yang, Yun Du, and Huiyu Zhou. Semantically contrastive learning for low-light image enhancement. In *AAAI*, pages 1555–1563, 2022.
- [4] 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 *IEEE CVPR*, pages 10561–10570, 2021.
- [5] Long Ma, Tengyu Ma, Risheng Liu, Xin Fan, and Zhongxuan Luo. Toward fast, flexible, and robust low-light image enhancement. In *IEEE CVPR*, 2022.
- [6] Yufei Wang, Renjie Wan, Wenhan Yang, Haoliang Li, Lap-Pui Chau, and Alex C. Kot. Low-light image enhancement with normalizing flow. In *AAAI*, pages 2604–2612, 2022.
- [7] Xiaogang Xu, Ruixing Wang, Chi-Wing Fu, and Jiaya Jia. Snr-aware low-light image enhancement. In *IEEE CVPR*, pages 17714–17724, 2022.