URWKV: Unified RWKV Model with Multi-state Perspective for Low-light Image Restoration (Supplementary Material)

Rui Xu[1,2], Yuzhen Niu[1,2]*, Yuezhou Li[1,2], Huangbiao Xu[1,2], Wenxi Liu[1,2], Yuzhong Chen[1,2]
[1]Fujian Key Laboratory of Network Computing and Intelligent Information Processing, College of Computer and Data Science, Fuzhou University, Fuzhou 350108, China
[2]Engineering Research Center of Big Data Intelligence, Ministry of Education, Fuzhou 350108, China
{xurui.ryan.chn, yuzhenniu, liyuezhou.cm, huangbiaoxu.chn}@gmail.com
wenxi.liu@hotmail.com, yzchen@fzu.edu.cn

In this supplementary material, we provide additional details and experimental results. First, we outline the framework details of URWKV in Section 1. Next, Section 2 offers descriptions of the datasets used and the training settings. In Section 3, we present further ablation studies to analyze the model from different perspectives. Section 4 provides more qualitative examples to illustrate the model's performance. Section 5 also evaluates the efficiency of models with different architecture. Finally, in Section 6, we discuss the limitations of the URWKV model and outline potential directions for future work.

1. Framework Details

The pipeline of our URWKV is outlined in the main paper. Here, we provide a more detailed description of the process for handling the input degraded image.

Given an input image $I \in \mathbb{R}^{3\times H\times W}$, we first use an input projection layer to extract shallow features $I_{in} \in \mathbb{R}^{C\times H\times W}$. These shallow features then pass through three encoder stages and three decoder stages, with an output projection layer generating the residual image $R \in \mathbb{R}^{3\times H\times W}$. Between the encoder and decoder stages, the state-aware selective fusion (SSF) module selectively fuses the contextual information from multiple encoder states into the decoder, by perceiving the states of the encoder stages. Ultimately, the output restored image can be expressed as $I' = I + R$.

Both the input and output projection layers consist of a 3×3 convolutional layer. In the encoder, each stage consists of N_1 URWKV blocks followed by a down-sampling layer, while each stage in the decoder contains N_2 URWKV blocks and an up-sampling layer. The down-sampling layer first doubles the number of channels through a 3×3 convolution, then halves the spatial resolution using bilinear down-sampling. In contrast, the up-sampling layer reduces the channel dimension by half via a 3×3 convolution and then doubles the spatial resolution using bilinear interpolation. Notably, the down-sampling operation in the first encoder stage and the up-sampling operation in the first decoder stage do not alter the number of channels.

2. Datasets and Training Details

Our URWKV model employs the same multi-scale progressive training strategy [1, 2] and the unified loss function [3] across all datasets. Following Retinexformer [1], all image pairs in the SID, SMID, SDSD-indoor, and SDSD-outdoor datasets are resized to 960×512. Furthermore, the RAW data are converted into low-light and normal-light RGB image pairs. A detailed description of each dataset utilized for comparison is provided below.

LOL. We evaluate models under both real-world (LOL-v2-real [4]) and synthetic (LOL-v2-syn [4]) low-light conditions, which pose a fundamental challenge in restoring both brightness and contrast while mitigating noise. The LOL-v2-real subset comprises 789 paired low-light and high-light images, captured in real-world scenarios by varying camera settings such as ISO and exposure time, with a resolution of 400×600. Among these, 689 pairs are designated for training, while 100 pairs are reserved for testing. In contrast, the LOL-v2-syn subset synthesizes low-light images from RAW data by modeling the illumination distribution of real low-light scenes. This subset features a resolution of 384×384 and includes 900 pairs for training and 100 pairs for testing.

SID and SMID. The SID [5] and SMID [5] datasets feature short- and long-exposure images with significant noise. Specifically, the SID subset consists of 2,697 paired short- and long-exposure images captured using the Sony α7S II

*Corresponding author.

camera. It encompasses both indoor and outdoor scenes under extremely low-light conditions, with illuminance levels ranging from 0.03 to 0.3 lux indoors and 0.2 to 5 lux outdoors. The dataset is partitioned into 2,099 pairs for training and 598 pairs for testing. In addition, the SMID subset comprises 20,809 paired short- and long-exposure images in RAW format. The dataset is partitioned into 15,763 pairs for training, with the remaining 5,046 pairs reserved for testing.

SDSD. The static version of the SDSD dataset, used in this work, is captured with a Canon EOS 6D Mark II camera equipped with an ND filter to regulate light exposure. It comprises both indoor and outdoor subsets [6], characterized by extremely low-brightness conditions. For SDSD-indoor, 1,655 image pairs are used for training and 308 for testing. For SDSD-outdoor, 2,650 image pairs are allocated for training and 500 for testing.

FiveK. The MIT-Adobe FiveK dataset [7] challenges models with the intricate task of color restoration in low-light and underexposed images. It [7] consists of 5,000 images in diverse lighting conditions. Each image is manually adjusted by five photographers (A-E). Following [1], we use the adjustments made by Expert C as the ground truth. It is split into 80% for training (4,500 pairs) and 20% for testing (500 pairs), with all images processed in sRGB output mode.

LOL-blur. The LOL-blur [8] dataset couples varying degrees of low-light and motion blur, which has often been underrepresented in other LLIE datasets. In particular, the LOL-blur dataset [8] contains 12,000 image pairs, each consisting of a low-blur and normal-sharp version. It is split into 10,200 pairs for training and 1,800 pairs for testing, with each image pair having a resolution of 1120×640.

3. More Ablation Studies

In this section, we conduct additional ablation studies to investigate the underlying design rationale and assess the model's scalability.

Additional analysis of LAN and SSF modules. LAN and SSF modules constitute essential components of the URWKV block. To gain a deeper understanding of their roles and contributions to the overall model performance, we conduct an ablation analysis on the SID and LOL-blur datasets. The SID dataset, which includes severe noise degradations in low-light conditions, and the LOL-blur dataset, which features coupled blur degradations, provide valuable insights.

Taking the results on the SID dataset as an example, as shown in Table 1a, the model without LAN and SSF performs the worst in terms of PSNR and SSIM. Notably, when only the SSF module is used (Table 1b), the improvements in PSNR and SSIM are limited. This may be due to the extreme low-light conditions, which pose significant challenges for the model. In such conditions, the model lacks tailored strategies for adaptive luminance adjustment. On the other hand, as shown in Table 1c, the gain from using only the LAN module is quite noticeable. Compared to the model in Table 1a, there is an improvement of 0.44 dB in PSNR and 0.003 in SSIM. This highlights that in complex environments, luminance, as one of the primary causes of other degradations, needs to be addressed specifically. Our final model, which combines both LAN and SSF, achieves a PSNR of 23.11 dB and an SSIM of 0.673, significantly outperforming other models.

Table 1. Ablation study of LAN and SSF modules on SID and LOL-blur datasets.

#	Models		SID		LOL-blur	
	LAN	SSF	PSNR ↑	SSIM ↑	PSNR ↑	SSIM ↑
(a)			21.75	0.658	24.92	0.844
(b)		✓	21.95	0.659	26.64	0.879
(c)	✓		22.19	0.661	25.95	0.873
(d)	✓	✓	23.11	0.673	27.27	0.890

Table 2. Impacts of EMA-based aggregation strategy on LOL-v2-real dataset. NAN indicates an out-of-memory condition. T represents the restoration stage, and C denotes the number of channels.

Aggregation Strategy	PSNR ↑	SSIM ↑	Complexity
Naive Attention [9]	NAN	NAN	$O(T \cdot N^2)$
Transposed Attention [10]	22.48	0.864	$O(T \cdot C^2)$
Naive Addition	21.99	0.858	$O(1)$
EMA (Ours)	23.11	0.874	$O(1)$

We also provide a visual example on the SID dataset for analysis. As shown in the first row of Fig. 1, the model without LAN exhibits noticeable blotchy artifacts, while the model without SSF fails to effectively suppress noise, resulting in a lack of smoothness. When both LAN and SSF are combined, our URWKV model achieves smoother noise suppression and produces sharper edge reconstruction. Additionally, our URWKV model not only demonstrates more accurate luminance enhancement (as shown in the second row of Fig. 1), but also delivers more precise color restoration (as shown in the third row of Fig. 1).

Aggregation strategy in SQ-Shift. The SQ-Shift employs an exponential moving average (EMA)-based approach to aggregate multiple intra-stage states. To evaluate its effectiveness, we conduct an ablation study by replacing EMA with alternative strategies, including additive (linear) aggregation and transformer-based (non-linear) aggregation. For the transformer-based approach, we explore two attention mechanisms: naive attention [9] and transposed attention [10]. The results presented in Table 2 highlight the advantages of the EMA approach. While linear aggre-

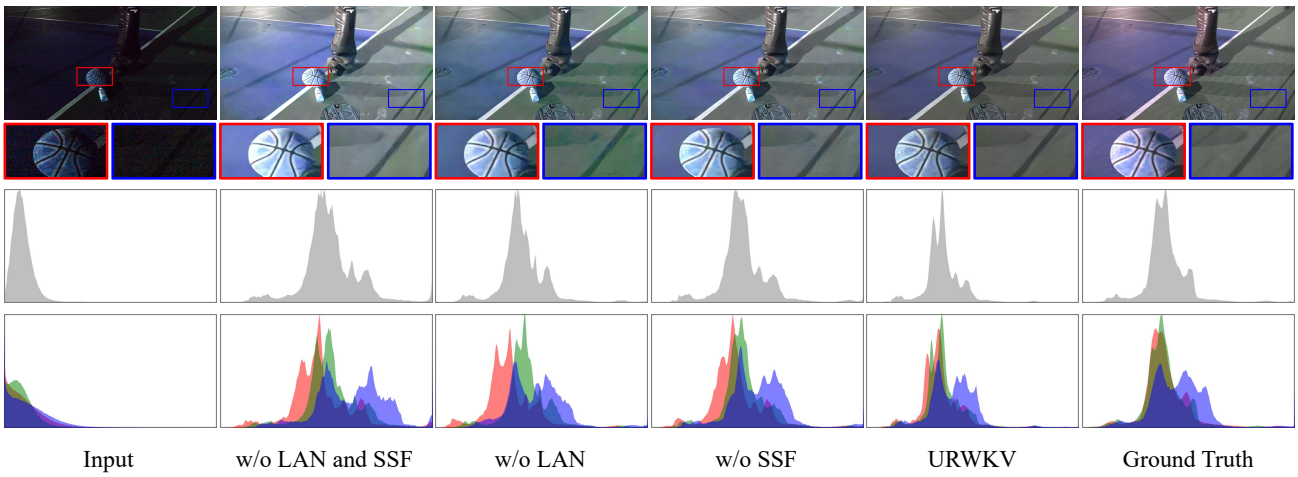


Figure 1. Visual analysis of LAN and SSF modules on the SID dataset. In particular, the second and third rows correspond to the luminance histograms and RGB histograms of the respective images, respectively.

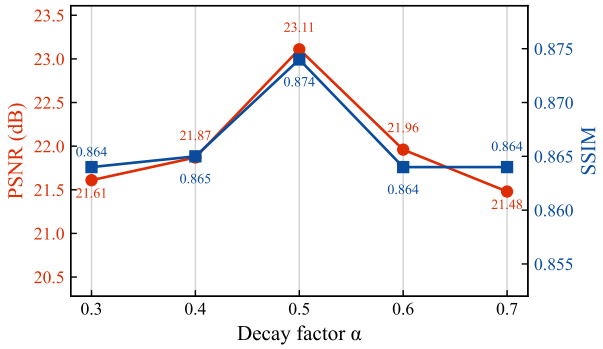


Figure 2. Ablation study on the decay factor α in EMA. Experiments are conducted on LOL-v2-real datatset.

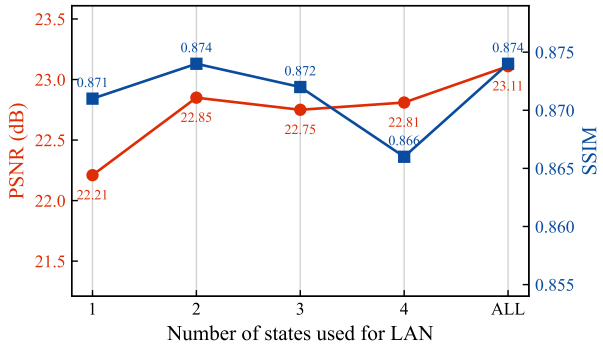


Figure 3. Ablation study on the number of states employed for LAN. Experiments are conducted on LOL-v2-real datatset.

gation is straightforward, it assigns equal weight to all past states, potentially diluting the significance of recent states that typically carry more relevant information. As shown in Table 2, the additive aggregation strategy results in suboptimal performance, yielding a PSNR of 21.99 and an SSIM of 0.858, respectively. On the other hand, transformer-based aggregation provides strong modeling capabilities but incurs substantial computational and memory overhead, making it an inefficient choice for aggregating intra-stage states in resource-constrained settings. In contrast, our model using EMA strikes an effective balance between efficiency and performance by dynamically adjusting the contribution of past and recent states using a tunable decay factor. This computationally lightweight approach introduces minimal overhead (linear complexity) while maintaining robustness, making it well-suited for the aggregation of intra-stage states in SQ-Shift.

Decay factor α in EMA. The decay factor α in EMA plays a pivotal role in regulating the integration of historical and current information within the model. For lower decay factors, the model's ability to retain useful historical information diminishes, resulting in weaker feature retention and a subsequent drop in performance. As shown in Fig. 2, when $\alpha = 0.3$ and $\alpha = 0.4$, the model yields suboptimal PSNR values of 21.61 dB and 21.87 dB, respectively. Conversely, higher decay factors lead to a decline in performance as the model becomes overly reliant on past information. As demonstrated in Fig. 2, the model achieves PSNR values of 21.96 dB and 21.48 dB for $\alpha = 0.6$ and $\alpha = 0.7$, respectively, both of which are suboptimal. These results reflect the negative impact of excessive historical influence. In contrast, the optimal performance is observed at $\alpha = 0.5$, which strikes an effective balance between historical context and current adaptability.

Number of states used for LAN. The LAN leverages inter-stage multiple states to dynamically adjust luminance, using an extended historical context across both encoder and decoder stages. To assess the number of states used for LAN on model performance, we conduct an ablation study using configurations with 1, 2, 3, and 4 states, as well as

and an "All" setting that integrates all available states. As illustrated in Fig. 3, increasing the number of states generally enhances performance by providing a richer historical context for luminance adjustment. For example, the PSNR improves from 22.21 dB with a single state to 22.85 dB with two states, highlighting the benefit of leveraging additional inter-stage information. However, as more states are added (e.g., three and four states), the performance gains begin to plateau, with only marginal improvements observed. This diminishing return may result from the inclusion of redundant or less relevant information from earlier stages, which can introduce noise or minor perturbations that disrupt the adjustment process. Notably, the "All" setting, which integrates the complete historical context, achieves the best results with a PSNR of 23.11 dB and an SSIM of 0.874. This finding indicates that, when sufficient and relevant historical information is fully utilized, the model can achieve a comprehensive understanding of luminance variations across stages.

Going deeper and wider. To enhance the capacity of our URWKV model, we explore increasing both the block number and the channel size, making the model deeper and wider, respectively. As shown in Table 3, our model demonstrates significant scalability, achieving substantial improvements in both configurations. For example, when the number of blocks in the encoder and decoder are increased from $N_1 = 3$, $N_2 = 2$ (Table 3a) to $N_1 = 4$, $N_2 = 3$ (Table 3b), the PSNR increases by 0.38 dB (from 27.27 dB to 27.65 dB), and SSIM improves by 0.002 (from 0.890 to 0.892). On the other hand, when the channel size is expanded from 32 (Table 3a) to 48 (Table 3c), the model exhibits a substantial performance boost with PSNR increasing from 27.27 dB to 27.37 dB and SSIM improving from 0.890 to 0.896. It is worth noting that even with our largest model (Table 3c), which has a parameter count comparable to the state-of-the-art models such as PDHAT [11], it achieves significant advantages with substantially fewer FLOPs.

4. More Qualitative Results

In this section, we present additional visual examples for various state-of-the-art models.

LOL. Visual comparisons on the LOL-v2-real dataset are shown in Fig. 4a. As illustrated in the first and second rows, our method excels in preserving details under scenes with significant brightness contrast, without introducing color shifts or oversaturated highlights in the illuminated areas. In contrast, other methods, such as SNR-Net [12], generate noticeable artifacts, while Retinexformer [1] produces large black blotches in the lighted regions. Additionally, Fig. 4b presents visual comparisons on the LOL-v2-syn dataset. Our method demonstrates outstanding detail preservation and color restoration in areas with pronounced brightness contrasts (e.g., the first and second rows). Moreover, it effectively avoids artifact generation, as evidenced by two clear examples in the third and fourth rows.

Table 3. Exploring deeper and wider URWKV models with LOL-blur dataset.

#	C	N_1	N_2	PSNR ↑	SSIM ↑	Params	FLOPs
(a)	32	3	2	27.27	0.890	2.25M	18.34G
(b)	32	4	3	27.65	0.892	2.83M	22.33G
(c)	48	3	2	27.37	0.896	5.05M	40.92G
(d)	48	4	3	28.11	0.903	6.36M	49.86G
LEDNet [8]	-	-	-	26.06	0.846	7.41M	38.57G
PDHAT [11]	-	-	-	26.71	0.879	7.83M	208.19G
MIRNet [13]	-	-	-	23.99	0.774	31.76M	785.00G
Restormer [2]	-	-	-	26.38	0.860	26.11M	140.99G
MambaIR [14]	-	-	-	26.28	0.848	4.30M	60.66G

SID and SMID. Enhancing brightness while simultaneously denoising poses a significant challenge for models on the SID and SMID datasets. Visual comparisons for these datasets are provided in Fig. 5 and Fig. 6, respectively. Our model demonstrates robust denoising capabilities while ensuring effective brightness enhancement and accurate color restoration, producing clean and sharp restored images. In contrast, other state-of-the-art models often retain considerable degradation patterns. For example, as shown in the first row of Fig. 5, the enhancement results from KinD [17] still exhibit a significant amount of residual noise, while FourLLIE [16] introduces noticeable color distortions. Similarly, in regions with stark brightness contrasts (e.g., the first-row example in Fig. 6), models such as MIRNet [13] and SNR-Net [12] fail to recover clear edges and instead produce noisy outputs.

SDSD. Visual examples on the SDSD-indoor and SDSD-outdoor datasets are presented in Fig. 7 and Fig. 8, respectively. As shown in the first row of Fig. 7, the LOL-deblur models, LEDNet [8] and PDHAT [11], introduce severe artifacts on white wall surfaces. Similarly, LLFormer [10] struggles with detail reconstruction and color restoration, as demonstrated in the second and third rows of Fig. 7. Furthermore, as illustrated in the first-row example of Fig. 8, SNR-Net [12] introduces noticeable block artifacts, a common drawback of transformer-based models when attempting to establish long-range relationships. In contrast, our URWKV model excels in these scenarios, delivering superior brightness enhancement while preserving details and effectively avoiding introducing artifact.

FiveK. Fig. 9 provides several visual examples from the FiveK dataset. On this dataset, SNR-Net [12] introduces prominent block artifacts, as seen in the first-row example. Retinexformer [1], on the other hand, suffers from severe color distortion, particularly in the sixth and seventh rows. The unified model MIRNet [13] exhibits suboptimal performance in detail reconstruction, as demonstrated in the

second-row example. Additionally, it leaves noticeable artifacts when restoring regions such as the sailboat, sky, or lake, as evidenced in the third, fifth, and sixth rows. In contrast, our model consistently delivers superior restoration across various scenes, showcasing exceptional aesthetic enhancement and structural reconstruction capabilities.

LOL-blur. The LOL-blur dataset introduces additional challenges by coupling complex blur degradation with low-light conditions, posing significant difficulties for all models. Several visual comparisons are presented in Fig. 10. Retinexformer [1], as a state-of-the-art LLIE model, lacks an explicit decoupling strategy. While it achieves some degree of brightness enhancement, its ability to restore blurred regions is minimal. Similar limitations can also be observed in unified models, such as MIRNet [13] and MambaIR [14]. In contrast, our model outperforms even LLIE-deblur models like LEDNet [8] in this challenging scenario. As shown in the first-row example, our model produces sharper edge restoration and achieves brightness enhancement that more closely matches the ground truth.

5. Efficiency Analysis

Smaller models (such as FourLLIE and Retinexformer) generally struggle with complex degradations, especially in challenging low-light scenarios coupled with heavy noise [5] or motion blur [8] (as visual examples depicted in Fig. 5 and Fig. 10) . Although larger models like MIRNet and MambaIR deliver more balanced performance, their extensive parameter counts and computational demands (31.76M/785.00G and 4.30M/60.66G) make them less suited for real-time deployment. In contrast, URWKV provides a robust solution with fewer parameters and lower FLOPs (**2.25M/18.34G**).

Furthermore, we compare the average inference speed of URWKV and MambaIR on the LOL-v2-real dataset, which are built on the advanced RWKV and Mamba architectures, respectively. URWKV achieves an inference time of 0.147s, while MambaIR takes 0.378s. This further highlights the ability of our approach to strike an optimal balance between performance and efficiency.

6. Limitation and Future Direction

In this work, we have tackled the challenges posed by dynamic coupled degradation in existing models and proposed a tailored solution from multi-state perspective. The proposed URWKV model demonstrates exceptional performance across eight datasets, with significantly fewer parameters and reduced computational cost. However, there remains potential for further improvement in restoring more extreme coupled degradations, such as severe noise and intense blur.

From an architectural standpoint, the potential of URWKV has yet to be fully explored. For example, integrating URWKV with Transformer-based architectures could enhance the model's ability to capture both local and global relationships across spatial and state domains, ultimately improving its performance in more complex degradation scenarios.

References

[1] Yuanhao Cai, Hao Bian, Jing Lin, Haoqian Wang, Radu Timofte, and Yulun Zhang. Retinexformer: One-stage retinex-based transformer for low-light image enhancement. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 12504–12513, 2023. 1, 2, 4, 5, 7, 9, 10

[2] Syed Waqas Zamir, Aditya Arora, Salman Khan, Munawar Hayat, Fahad Shahbaz Khan, and Ming-Hsuan Yang. Restormer: Efficient transformer for high-resolution image restoration. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 5728–5739, 2022. 1, 4

[3] Rui Xu, Yuezhou Li, Yuzhen Niu, Huangbiao Xu, Yuzhong Chen, and Tiesong Zhao. Bilateral interaction for local-global collaborative perception in low-light image enhancement. *IEEE Transactions on Multimedia*, 2024. 1

[4] Wenhan Yang, Wenjing Wang, Haofeng Huang, Shiqi Wang, and Jiaying Liu. Sparse gradient regularized deep retinex network for robust low-light image enhancement. *IEEE Transactions on Image Processing*, 30:2072–2086, 2021. 1

[5] Chen Chen, Qifeng Chen, Jia Xu, and Vladlen Koltun. Learning to see in the dark. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 3291–3300, 2018. 1, 5, 8

[6] Ruixing Wang, Xiaogang Xu, Chi-Wing Fu, Jiangbo Lu, Bei Yu, and Jiaya Jia. Seeing dynamic scene in the dark: A high-quality video dataset with mechatronic alignment. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 9700–9709, 2021. 2, 9

[7] Vladimir Bychkovsky, Sylvain Paris, Eric Chan, and Frédo Durand. Learning photographic global tonal adjustment with a database of input/output image pairs. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 97–104, 2011. 2

[8] Shangchen Zhou, Chongyi Li, and Chen Change Loy. LEDNet: Joint low-light enhancement and deblurring in the dark. In *Proceddings of the European Conference on Computer Vision*, pages 573–589, 2022. 2, 4, 5, 9, 10

[9] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, and Illia Polosukhin. Attention is all you need. In *Proceedings of the Advances in Neural Information Processing Systems*, pages 5998–6008, 2017. 2

[10] Tao Wang, Kaihao Zhang, Tianrun Shen, Wenhan Luo, Bjorn Stenger, and Tong Lu. Ultra-high-definition low-light image enhancement: A benchmark and transformer-based method. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 37, pages 2654–2662, 2023. 2, 4, 8, 9

[11] Yuezhou Li, Rui Xu, Yuzhen Niu, Wenzhong Guo, and Tiesong Zhao. Perceptual decoupling with heterogeneous

auxiliary tasks for joint low-light image enhancement and deblurring. *IEEE Transactions on Multimedia*, pages 6663–6675, 2024. 4, 9

[12] Xiaogang Xu, Ruixing Wang, Chi-Wing Fu, and Jiaya Jia. SNR-aware low-light image enhancement. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 17714–17724, 2022. 4, 7, 8, 9

[13] Syed Waqas Zamir, Aditya Arora, Salman Khan, Munawar Hayat, Fahad Shahbaz Khan, Ming-Hsuan Yang, and Ling Shao. Learning enriched features for real image restoration and enhancement. In *Proceedings of the European Conference on Computer Vision*, pages 492–511, 2020. 4, 5, 8, 9, 10

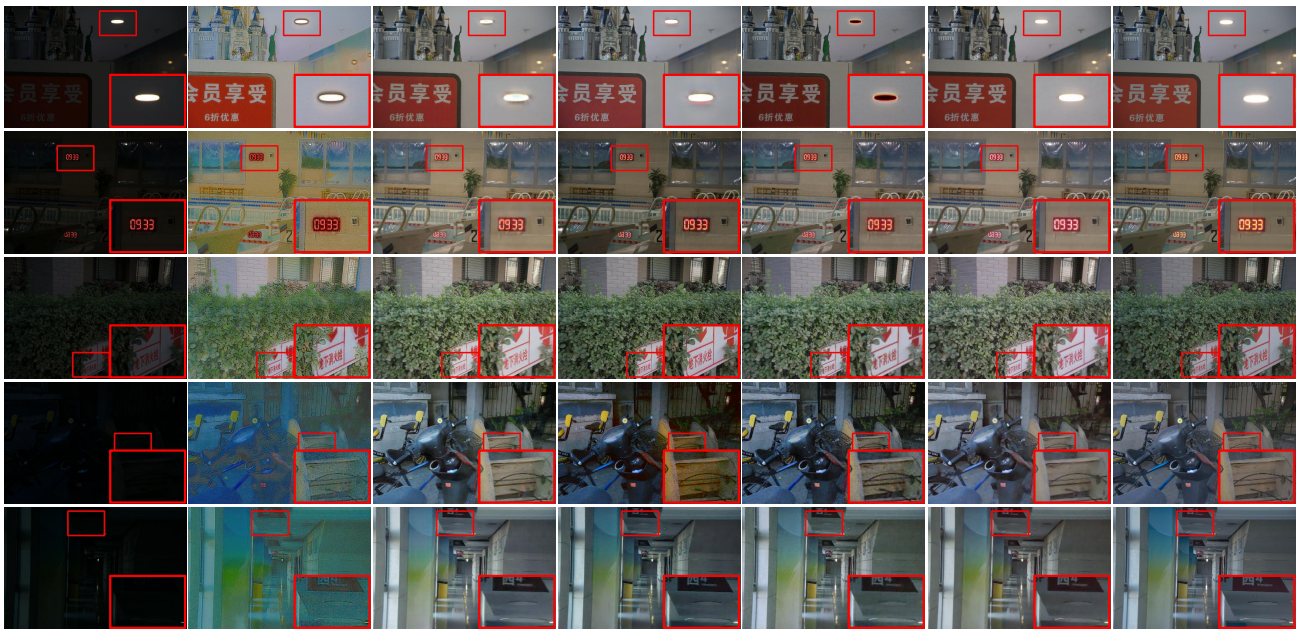
[14] Hang Guo, Jinmin Li, Tao Dai, Zhihao Ouyang, Xudong Ren, and Shu-Tao Xia. MambaIR: A simple baseline for image restoration with state-space model. In *Proceedings of the European Conference on Computer Vision*, pages 222–241, 2025. 4, 5, 10

[15] Chen Wei, Wenjing Wang, Wenhan Yang, and Jiaying Liu. Deep retinex decomposition for low-light enhancement. In *Proceedings of the British Machine Vision Conference*, 2018. 7

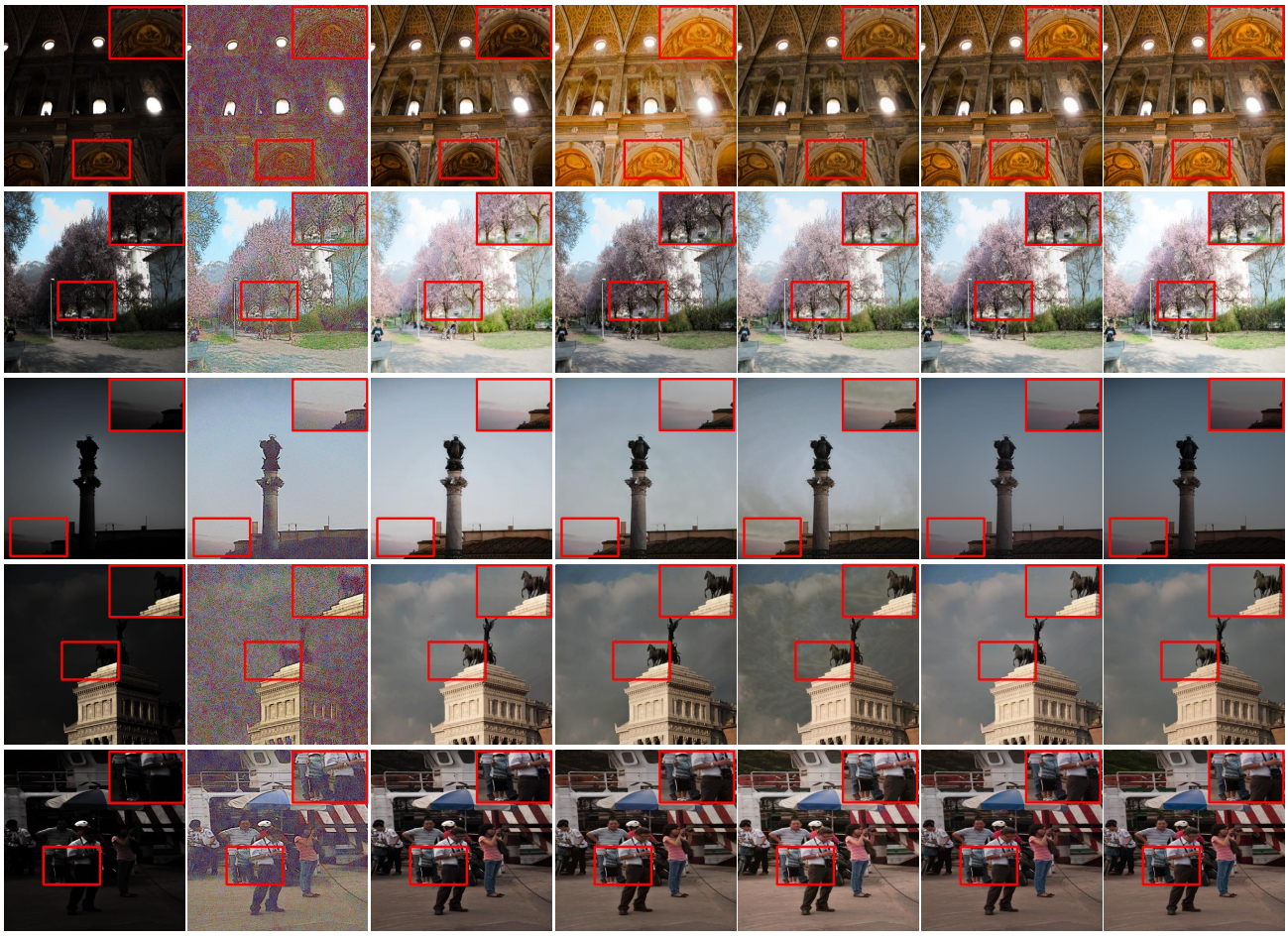
[16] Chenxi Wang, Hongjun Wu, and Zhi Jin. FourLLIE: Boosting low-light image enhancement by fourier frequency information. In *Proceedings of the ACM International Conference on Multimedia*, pages 7459–7469, 2023. 4, 7, 8, 9

[17] Yonghua Zhang, Jiawan Zhang, and Xiaojie Guo. Kindling the darkness: A practical low-light image enhancer. In *Proceedings of the ACM International Conference on Multimedia*, pages 1632–1640, 2019. 4, 8

[18] Jiesong Bai, Yuhao Yin, and Qiyuan He. Retinexmamba: Retinex-based mamba for low-light image enhancement. *arXiv preprint arXiv:2405.03349*, 2024. 9



(a) LOL-v2-real



(b) LOL-v2-syn

Figure 4. Visual examples on LOL-v2-real and LOL-v2-syn datasets, comparing RetinexNet [15], SNR-Net [12], FourLLIE [16], Retinexformer [1], and our URWKV.

Figure 5. Visual examples on SID [5] among KinD [17], SNR-Net [12], FourLLIE [16], LLFormer [10], and our URWKV.

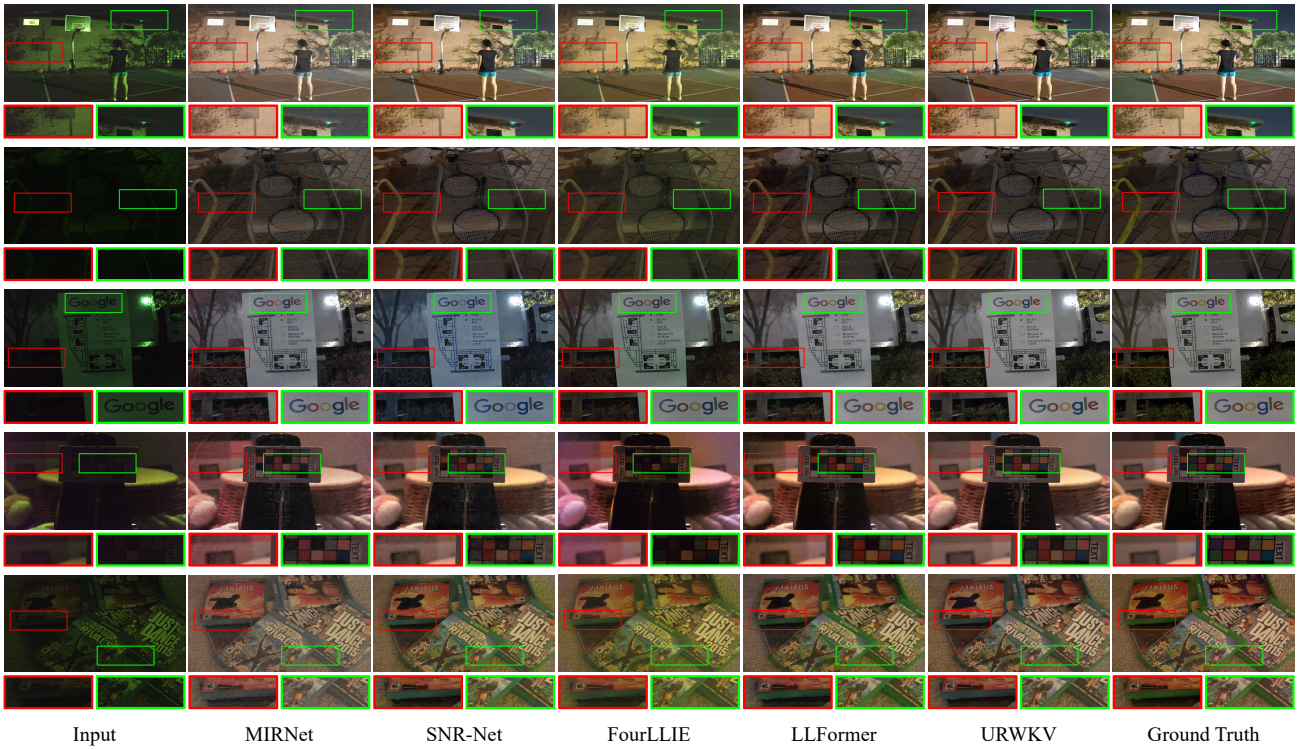


Figure 6. Visual examples on SMID [5] among MIRNet [13], SNR-Net [12], FourLLIE [16], LLFormer [10], and our URWKV.

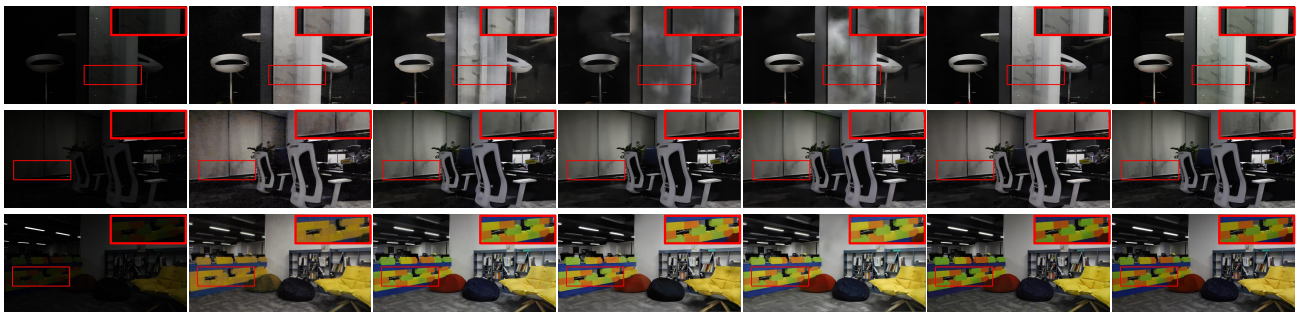


Figure 7. Visual examples on SDSD-indoor [6] among LLFormer [10], RetinexMamba [18], LEDNet [8], PDHAT [11], and our URWKV.

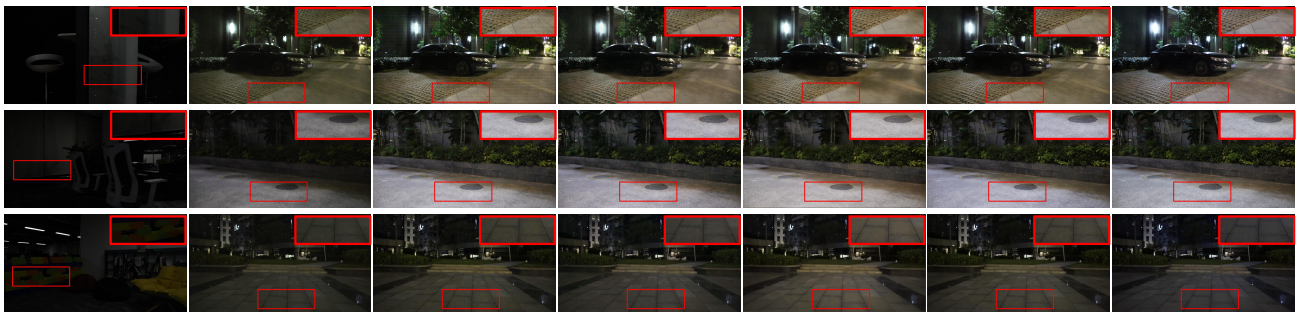


Figure 8. Visual examples on SDSD-outdoor [6] among FourLLIE [16], SNR-Net [12], LLFormer [10], LEDNet [8], and our URWKV.

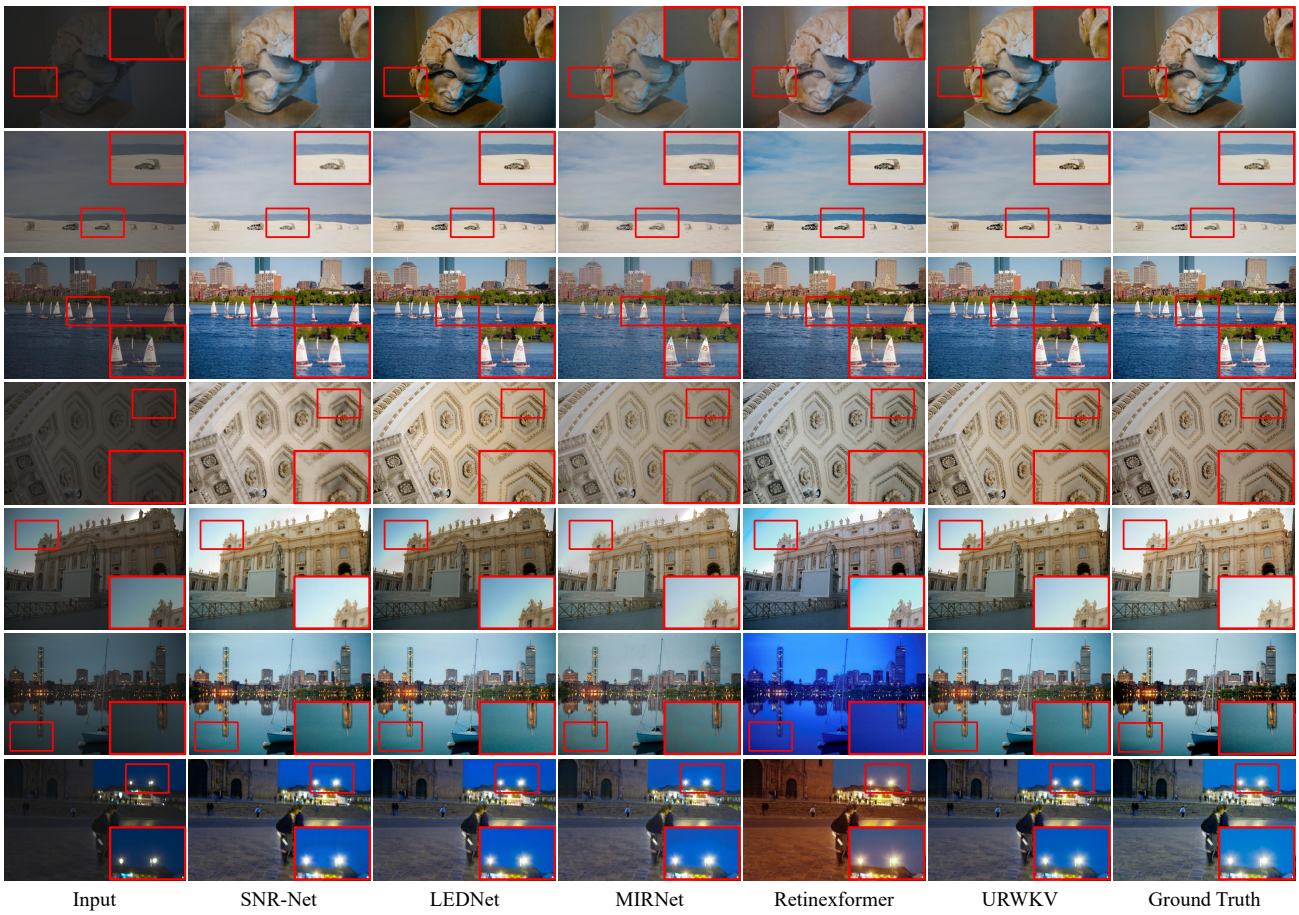


Figure 9. Visual examples on FiveK [6] among SNR-Net [12], LEDNet [8], MIRNet [13], Retinexformer [1], and our URWKV.

Figure 10. Visual examples on LOL-blur [8] among Retinexformer [1], LEDNet [8], MIRNet [13], MambaIR [14], and our URWKV.