PBDL-Challenge-IMAGCX for the Low-Light-srgb-enhancement track Technical Report

Xiang Chen Hao Li Jinshan Pan
Nanjing University of Science and Technology

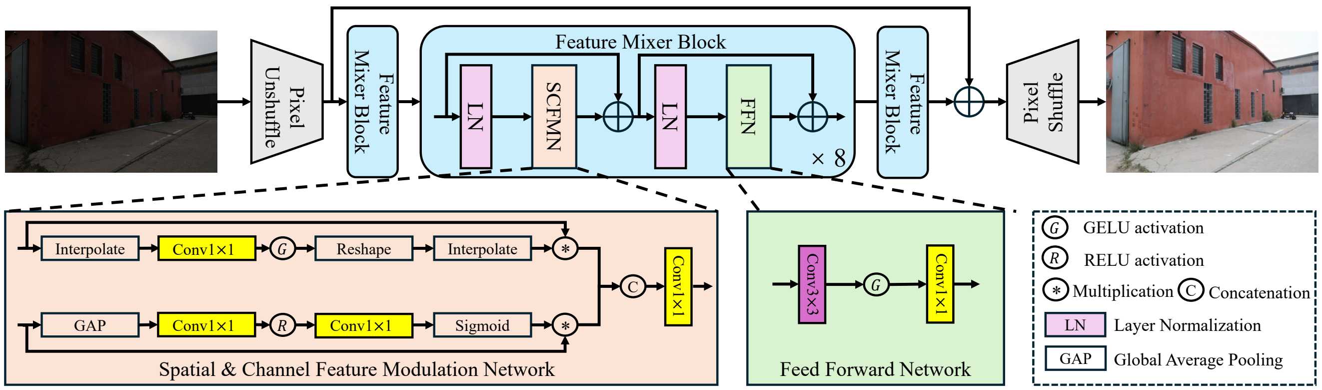

Figure 1. Overview of Team IMAGCX's technique for low-light-srgb-enhancement track.

In this report, we introduce the technical solution and training details of our team.

This competition adopts the low-light image enhancement dataset proposed by Fu *et al.* [1]. The training data contains 1,700 paired sRGB ultra-high-definition (UHD) 4K images. To solve UHD low-light image enhancement, several recent state-of-the-art methods have been proposed, for example LLFormer [4], UHDFour [2], and MixNet [5]. We first conduct cross-domain generalization analysis on these methods, and we find that MixNet can better generalize to unseen real images. Thus, MixNet [5] is employed as the network backbone for low-light image enhancement.

Figure 1 shows the overview of the network architecture. It aims to map an UHD low-light input image $x \in \mathbb{R}^{H\times W\times C}$ to its corresponding normal-clear version $y \in \mathbb{R}^{H\times W\times C}$, where H, W, and C represent height, width, and channel, respectively. To reduce computational complexity, it downsample the input to 1/4 of the original resolution by PixelUnshuffle. Subsequently, the shallow features go through multiple deep feature mixer blocks. Each feature mixer block mainly consists of a feature modulation network and a feed forward network. To better capture long-range pixel dependencies in UHD images, feature modulation network combines spatial and channel dimensions for joint feature modeling. Finally, we use PixelShuffle upsampling to reconstruct the final image.

We conduct model training in PyTorch framework on 8 NVIDIA GeForce RTX 4090 GPUs. Furthermore, we incorporate other public UHD low-light image enhancement datasets (UHD-LL [2] and UHD-LOL [4]) into the network training. Similar to [3], patches at the size of 2000×2000 are randomly cropped from the image pairs as training samples. The training data is augmented with random rotation and flipping. To optimize the network, we adopt L1 loss as the optimization objective, and we employ the Adam optimizer with a learning rate 2×10^{-4}. In total, we perform 600k iterations. During the testing phase, we perform full-resolution inference using one NVIDIA GeForce RTX 4090 GPU. Note that we employ a self-ensemble strategy to further improve performance. The code and model are released at `https://drive.google.com/file/d/11Yn6Q4gCjLxWllwApaWRJW5K5DRyylF4/view?usp=sharing`.

References

[1] Ying Fu, Yang Hong, Linwei Chen, and Shaodi You. Le-gan: Unsupervised low-light image enhancement network using attention module and identity invariant loss. *Knowledge-Based Systems*, 240:108010, 2022. 1

[2] Chongyi Li, Chun-Le Guo, Man Zhou, Zhexin Liang, Shangchen Zhou, Ruicheng Feng, and Chen Change Loy. Embedding fourier for ultra-high-definition low-light image enhancement. In *ICLR*, 2023. 1

[3] Xiaoning Liu, Zongwei Wu, Ao Li, Florin-Alexandru Vasluianu, Yulun Zhang, Shuhang Gu, Le Zhang, Ce Zhu,

Radu Timofte, Zhi Jin, et al. Ntire 2024 challenge on low light image enhancement: Methods and results. *arXiv preprint arXiv:2404.14248*, 2024. 1

[4] 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 *AAAI*, pages 2654–2662, 2023. 1

[5] Chen Wu, Zhuoran Zheng, Xiuyi Jia, and Wenqi Ren. Mixnet: Towards effective and efficient uhd low-light image enhancement. *arXiv preprint arXiv:2401.10666*, 2024. 1