PBDL Challenge on Low Light SRGB Image Enhancement

Chuanlong Xie[1], Hongming Chen[1], Mingrui Li[2], Tianchen Deng[3], Jingwei Huang[4], Yufeng Li[1]
[1]Shenyang Aerospace University
[2]Dalian University of Technology
[3]Shanghai Jiao Tong University
[4]University of Electronic Science and Technology of China

Download link. The source code and pre-trained model are available at https://drive.google.com/file/d/1tPxeQBzI_ELlAmmoA70j2Xh98MxFFV4H/view?usp=sharing.

Network architecture. Figure 1 illustrates the overall architecture of our method. Specifically, the input x is first reshaped to feature tensor via PixelUnshuffle ($4\times\downarrow$) to preserve original information, and then fed to 8 feature extraction modules. Finally, the output feature y is reshaped to the original height and width of input x via Pixelshuffle ($4\times\uparrow$). The feature extraction module mainly contains a feature rearrangement block (FRB), a feature enhancement block (FEB), and a feed-forward network (FFN). Here, FRB adopts MLP-based tensor dimensional transformations [2], while FEB employs CNN-based local operators [1]. The overall process can be represented as follows:

$$\begin{aligned} F_1 &= \mathrm{Conv}\left[\mathrm{FRB}\left(\mathrm{LN}\left(F_0\right)\right); \mathrm{FEB}\left(\mathrm{LN}\left(F_0\right)\right)\right] + F_0, \\ F_2 &= \mathrm{FFN}\left(\mathrm{LN}\left(F_1\right)\right) + F_1, \end{aligned} \tag{1}$$

where F_0 denote the input features, F_1 denote the intermediate features and F_2 denote the output features. LN refers to the layer normalization.

Training strategy. To supervise the training process, we employ the L1 loss as the objective function. We conduct model training on 4 NVIDIA TESLA V100s with 32GB memory. In total, we perform 500 epochs of training. During the training, we adopt the Adam optimizer with a learning rate of 2×10^{-4}. The patch size is set to be 768×768 pixels and the batch size is set to be 16. To augment the training data, we apply random horizontal and vertical flips. For testing images, we use one NVIDIA GeForce RTX 4090 GPU with 24GB memory.

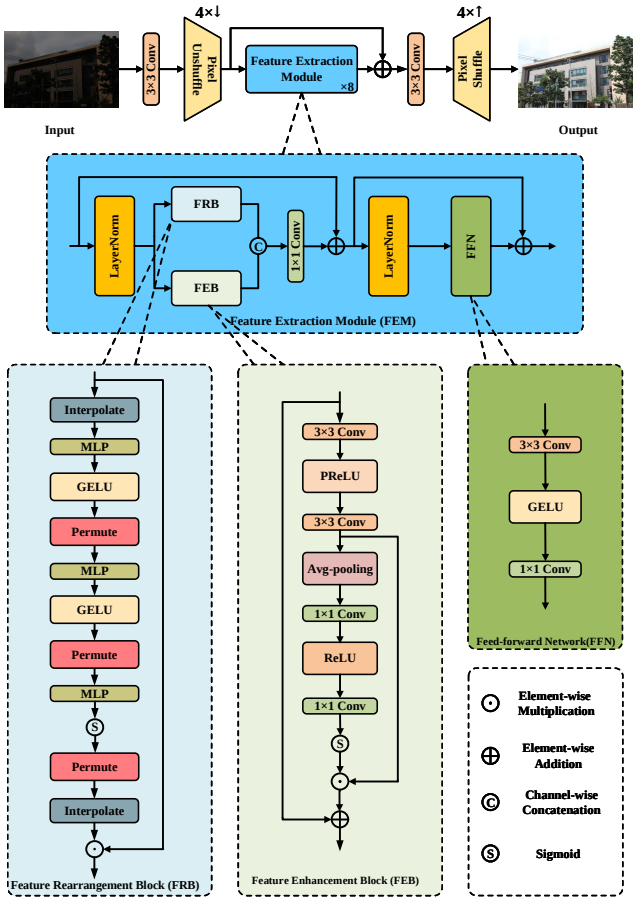

Figure 1. The network architecture of team SuperGo.

References

[1] Syed Waqas Zamir, Aditya Arora, Salman Khan, Munawar Hayat, Fahad Shahbaz Khan, Ming-Hsuan Yang, and Ling Shao. Multi-stage progressive image restoration. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 14821–14831, 2021. 1

[2] Yunliang Zhuang, Zhuoran Zheng, Yuang Zhang, Lei Lyu, Xiuyi Jia, and Chen Lyu. Dimensional transformation mixer for ultra-high-definition industrial camera dehazing. *IEEE Transactions on Industrial Informatics*, 2023. 1