NTIRE 2023 Efficient SR Challenge Factsheet
-Efficient Feature Distillation Network-

Kang Hu, Jingpeng Shi
Anhui University, Hefei, China
hukangyy@gmail.com,jinpeeeng.s@gmail.com

1. Team Details

- Team name: **FRL Team 04**
- Team leader name: **Kang Hu**
- Team leader address, phone number, and email:
  - address: **Anhui University, Hefei, China**
  - phone number: **+86 180 5620 0045**
  - email: **hukangyy@gmail.com**
- Rest of the team members: **Jinpeng Shi (advisor)**
- Team website URL (if any): github.com/Fried-Rice-Lab/FriedRiceLab
- Affiliation: **Anhui University**
- Affiliation of the team and/or team members with NTIRE 2023 sponsors (check the workshop website): **N/A**
- User names and entries on the NTIRE 2023 Codalab competitions (development/validation and testing phases):
  - user name: **Y__Y**
  - development entries: **6**
  - validation entries: **1**
- Best scoring entries of the team during development/validation phase:

PSNR	SSIM	Runtime	Params	Extra Data
28.97 (20)	0.83 (22)	0.06 (19)	173478.00 (10)	1.00 (1)

- Link to the codes/executables of the solution(s): github.com/Kkaif/NTIRE2023_ESR

Fried Rice Lab (FRL) is organized by students from Anhui University who are interested in image restoration. FRL is dedicated to proposing clean and efficient image restoration solutions and contributing to the image restoration open source community. **FRL Team 04**, led by Kang Hu and advised by Jinpeng Shi, was among the teams that FRL sent to compete in the NTIRE 2023 ESR competition.

2. Method Details

We follow the general structure of RFDN [1], which is divided into 4 stages: shallow feature extraction, deep feature extraction, multi-layer feature fusion and reconstruction. The RFDN is effectively derived from the residual feature distillation network(RFDB). We believe that the RFDB module is similar to the deepwise separable convolution [2] in that it operates pointwise convolution and deepwise convolution separately, using 3*3 convolution of residuals for intra-channel information interaction and 1*1 convolution for inter-channel information interaction, respectively. We think the part of deepwise can be replaced by the module of pixel mixing (PM) without parameters to make it more efficient. In addition, BN layer [3]is added to make it more generalizable. Efficient feature distillation network (EFDN) is proposed as shown in Figure 1. Specifically, in the first stage, shallow feature extraction consists of linear mapping and deepwise convolution from the input image space to a higher dimensional feature space. Then stacked efficient pixel mixing Blocks (EPMB)(Figure 1) are constructed for deep feature extraction, progressively refining the extracted features.The features generated by each EPMB are fused at the end of the backbone along the channel dimension.

2.1. Pixel Mixing Block

Pooling preserves the main features while reducing the number of parameters and computations [4]. In the process of information filtering, Pooling tends to retain the high-level semantic features while losing some of them. To avoid the loss of low-level visual features while reducing parameters and computation, and to enhance the network

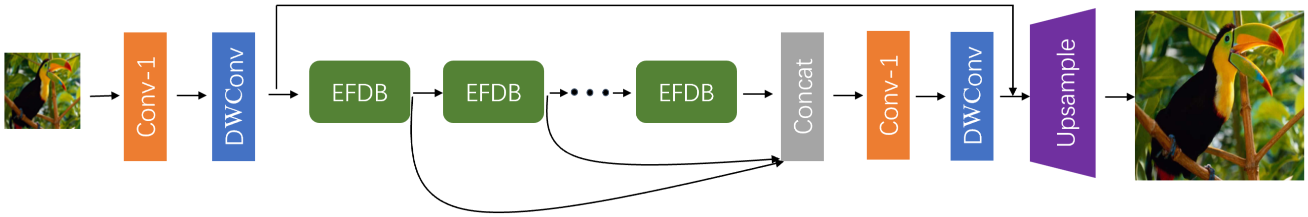

Figure 1. The architecture of Efficient feature distillation network (EFDN)

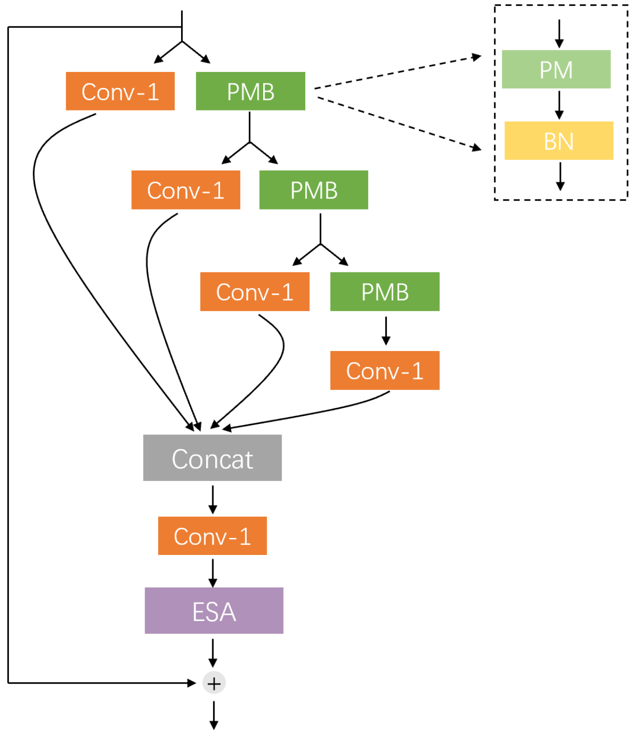

Figure 2. The architecture of the proposed efficient distillation network (EFDB).

generalization performance, we propose the module PM, which also has no parameters. PM divides the input into five groups equally, and moves the edge features to the opposite position in the order of left, right, top, and bottom for each of the first four groups, and keeps a copy of the features mixed with the relatively moved features. We describe this process as Pixel Mixing.

2.2. Training Details

Training was performed on DIV2K [5] and Flickr2K [6] images. HR patches of size 256 × 256 were randomly cropped from the HR images and the mini-batch size was set to 128. The model was trained with the ADAM optimizer, where $\beta_1 = 0.9$ and $\beta_2 = 0.9999$. The initial learning rate was set to 5×10^{-4} with cosine learning rate decay. The L2 loss was used for ab initio training and the number of iterations The model was implemented using Pytorch 1.10.1 and trained on 2 GeForce RTX 3090 GPUs.

2.3. Experimental Results

The experimental result is shown in Tbale 1. FLOPs and Activation are tested on an LR image of size 256×256. PNSR[val] is tested on DIV2k validation dataset, while PSNR[test] is calculated on a combination of DIV2K and LSDIR test data. The runtime is is evaluated on DIV2K and LSDIR test datasets using a single GeForce RTX 3090 GPU.

3. Other details

- Planned submission of a solution(s) description paper at NTIRE 2023 workshop.

Table 1. Result for NTIRE2023 ESR Challenge. FLOPs and Activation are tested on an LR image of size 256×256. The runtime is is averaged on DIV2K and LSDIR test datasets using a single NVIDIA RTX 3090 GPU.

PSNR[val]	PSNR[test]	Params[M]	FLOPs[G]	GPU Mem.[M]	Activation[M]	Average Runtime[ms]
28.97	27.02	0.1734	10.60	1266.92	187.32	55.16

We are not planning to submit the solution description paper to NTIRE2023 workshop, since it has been submitted to other conference.

- General comments and impressions of the NTIRE 2023 challenge.

  The organizers provided detailed processes and instructions and we appreciate the effort they spent!

References

[1] Liu J, Tang J, Wu G. Residual feature distillation network for lightweight image super-resolution[C]//Computer Vision–ECCV 2020 Workshops: Glasgow, UK, August 23–28, 2020, Proceedings, Part III 16. Springer International Publishing, 2020: 41-55. 1

[2] Chollet F. Xception: Deep learning with depthwise separable convolutions[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2017: 1251-1258. 1

[3] Ioffe S, Szegedy C. Batch normalization: Accelerating deep network training by reducing internal covariate shift[C]//International conference on machine learning. pmlr, 2015: 448-456. 1

[4] Gholamalinezhad H, Khosravi H. Pooling methods in deep neural networks, a review[J]. arXiv preprint arXiv:2009.07485, 2020. 1

[5] Eirikur Agustsson and Radu Timofte. Ntire 2017 challenge on single image super-resolution: Dataset and study. In CVPR workshops, pages 126–135, 2017. 2

[6] Radu Timofte, Eirikur Agustsson, Luc Van Gool, Ming Hsuan Yang, and Lei Zhang. Ntire 2017 challenge on single image super-resolution: Methods and results. In CVPR workshops, pages 114–125, 2017. 2