Lightweight Residual Feature Distillation Network for Efficient Image Super-Resolution

Jianbin Zheng
South China University of Technology
Guangzhou, Guangdong, China
jabir.zheng@outlook.com

1. Team details

- Team name
  Alpha
- Team leader name
  Jianbin Zheng
- Team leader address
  South China University of Technology, Guangzhou, Guangdong, China;
  phone number
  +86-13025191543;
  email
  2327396173@qq.com
- Rest of the team members
  only one member, *i.e.*, Jianbin Zheng
- Team website URL (if any): N/A
- Affiliation
  South China University of Technology
- 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: Jabir
  development phases entries: 7
  testing phases entries: 3
- Best scoring entries of the team during development/validation phase
  Entries 7 on the NTIRE 2023 Co- dalab competitions.
- Link to the codes/executables of the solution(s): https://github.com/jabir-zheng/LRFDN

2. Method details

General method description. We propose a Lightweight Residual Feature Distillation Network for Efficient Image Super-Resolution as shown in Fig. 1. Our work is mainly inspired by RFDN [2], MAFFSRN [4]. Following the overall architecture of RFDN, LRFDN consists of four stages including the shallow feature extraction, deep feature extraction, and reconstruction. To further reduce the parameters and computational complexity of the original RFDB module, the number of channels of layered distillation is effectively compressed. These distillation features are extracted by three shared 1×1 and one 3×3 convolutional filters. The design of the LRFDB block is shown in Fig. 2 (a).

Inspired by the residual block proposed in the RFANet [3], MAFFSRN [4] introduced an enhanced fast spatial attention module (EFSA). It aims to realize spatial attention weighting to make the features more concentrated in some desired regions, so that more representative features can be obtained. The design of the EFSA module is shown in Fig. 2 (c). Using the blocks above, the proposed model can better extract and integrate compact contextual information with fewer parameters.

Furthermore, it has been found that the channel-wise feature rescaling is effective for shallow SR models to boost reconstruction accuracy. Therefore, a channel weighting layer is involved in each LRFDB for modelling channel-wise relationships to utilize inter-dependencies among channels with slightly additional cost. Additionaly, we use GeLU activation function to replace LeakyReLU in RFDN.

Training strategy. The proposed LRFDN has four LRFDBs, in which the number of feature channels is set to 64. During training, DIV2K [1] and Flickr2K datasets are used for the whole process. The LRFDN model is trained from scratch with only one stage. HR patches of size 640×640 are randomly cropped from HR images, random horizontal flip, vertical flip and rotation are introduced

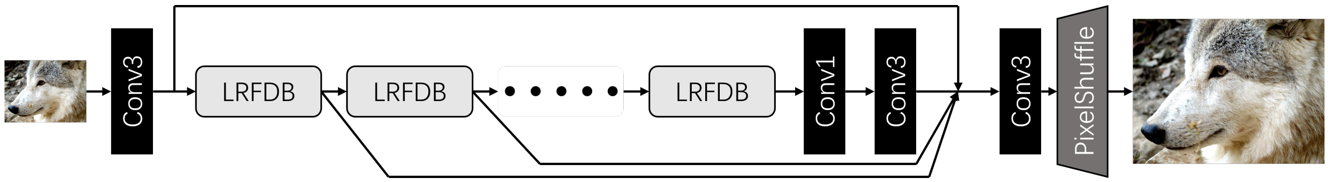

Figure 1. Framework of LRFDN.

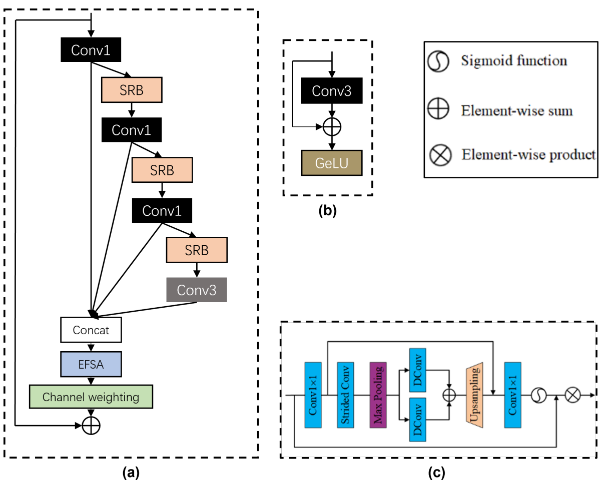

Figure 2. (a) Structure of IRFDN. (b) Structure of SRB. (c) Structure of EFSA.

as the data augmentation, and the mini-batch size is set to 64. We use Adam optimizer with $\beta_1 = 0.9$, $\beta_2 = 0.999$ to train our LRFDN model by minimizing L1 loss function. The base leaning rate is set to 5×10^{-4} equipped with cosine learning rate decay and 3000 warm up steps. The total number of epochs is 1000. We also maintain an exponential moving average (EMA) of LRFDN weights over training with a decay of 0.9999, which improves the PSNR by about 0.01dB on the DIV2K validation set.

Experimental results. As shown in the Tab. 1, we compared the LRFDN with the baseline RFDN on the DIV2K validation set, which was tested at A100 40G.

Model	PSNR(dB)	Time(ms)	Params(M)	FLOPs(G)	Acts(M)	Mem(M)	Conv
RFDN	29.04	20.91	0.433	27.10	112.03	788.13	64
LRFDN	28.81	19.26	0.198	11.23	86.18	768.76	56

Table 1. Comparison of RFDN and LRFDN

References

[1] Eirikur Agustsson and Radu Timofte. Ntire 2017 challenge on single image super-resolution: Dataset and study. *Computer Vision and Pattern Recognition*, 2017. 1

[2] Jie Liu, Jie Tang, and Gangshan Wu. Residual feature distillation network for lightweight image super-resolution. *Cornell University - arXiv*, 2020. 1

[3] Jie Liu, Wenjie Zhang, Yuting Tang, Jie Tang, and Gangshan Wu. Residual feature aggregation network for image super-resolution. *Computer Vision and Pattern Recognition*, 2020. 1

[4] Abdul Muqeet, Jiwon Hwang, Subin Yang, Jung Heum Kang, Yongwoo Kim, and Sung-Ho Bae. Multi-attention based ultra lightweight image super-resolution. *Lecture Notes in Computer Science*, 2020. 1