NTIRE 2024 Image Super-Resolution ($\times 4$) Challenge Factsheet
-Image Super-Resolution Reconstruction Using RepRLFN and HAT-

Weijun Yuan
Jinan University
Guangzhou, China
ywei jun@stu2022.jnu.edu.cn

Zhan Li
Jinan University
Guangzhou, China
lizhan@jnu.edu.cn

Wenqin Kuang
Jinan University
Guangzhou, China
manhing@stu2021.jnu.edu.cn

Ruijin Guan
Jinan University
Guangzhou, China
guanruijin@stu2021.jnu.edu.cn

Ruting Deng
Jinan University
Guangzhou, China
routine@stu2022.jnu.edu.cn

1. Network

Inspired by RLFN [4] and RepRFN [2], we proposed RepRLFN based on RLFN [4] using structural re-parameterization technology [2, 3], as shown in Fig 1. HAT [1] was also selected to train on the training set, which shows excellent performance in super-resolution tasks. The final results consist of the fusion of RepRLFN and HAT [1].

RepRLFN has the same structure as RepRFN [2], the difference is that RepRLFN replaces RepRFBs in RepRFN [2] with RepRLFBs. RepBlock is the main component of RepRLFB, which employs multiple parallel branch structures to extract the features of different receptive fields and modes to improve performance. At the same time, the structural re-parameterization technology is leveraged to decouple the training and inference phases to avoid the problem that computational complexity increases caused by the introduction of multi-branch.

The overall structure of HAT [1] consists of three parts, including shallow feature, deep feature extraction, and image reconstruction. HAT [1] combines both channel attention and window-based self-attention schemes to activate more pixels for super-resolution reconstruction.

2. Training strategy

For training RepRLFN, we used a large combination training dataset, which is composed of DIV2K, Flicker2K, and the first 10k images of LSDIR. HR images were randomly cropped into 480×480 patches, and LR images were cropped accordingly. Random horizontal/vertical flipping and RGB channel shuffling were adopted for data augmentation. The model was first trained with an L1 loss and then fine-tuned with an MSE loss using an Adam optimizer.

For training HAT [1], DIV2K and Flickr2K were used. The L1 loss is adopted with the learning rate of 1×10^{-5}. Data augmentation includes the horizontal flip and the rotation at 90 degrees. The batch size is 2, and the patch size is 64.

3. Testing description

In the testing phase, a test-time data ensemble strategy was adopted to improve the performance. In addition, the results of RepRLFN and HAT [1] are performed with a weighted fusion to obtain the final output.

4. Team details

- Team name: JNU_620
- Team leader name: Weijun Yuan
- Team leader address, phone number, and email: Jinan University, Guangzhou, 17817975121, yweijun@stu2022.jnu.edu.cn
- Rest of the team members: Zhan Li, Wenqin Kuang, Ruijin Guan, RutingDeng
- Affiliation: Jinan University
- User names and entries on the NTIRE 2024 Codalab competitions: Wedream
- Best scoring entries of the team during the development/validation phase: 30.7047,0.84 / 30.4325, 0.85
- Link to the codes/executables of the solution(s): https://github.com/Wedream-wj/NTIRE2024_ImageSR_x4

References

[1] Xiangyu Chen, Xintao Wang, Jiantao Zhou, Yu Qiao, and Chao Dong. Activating more pixels in image super-resolution

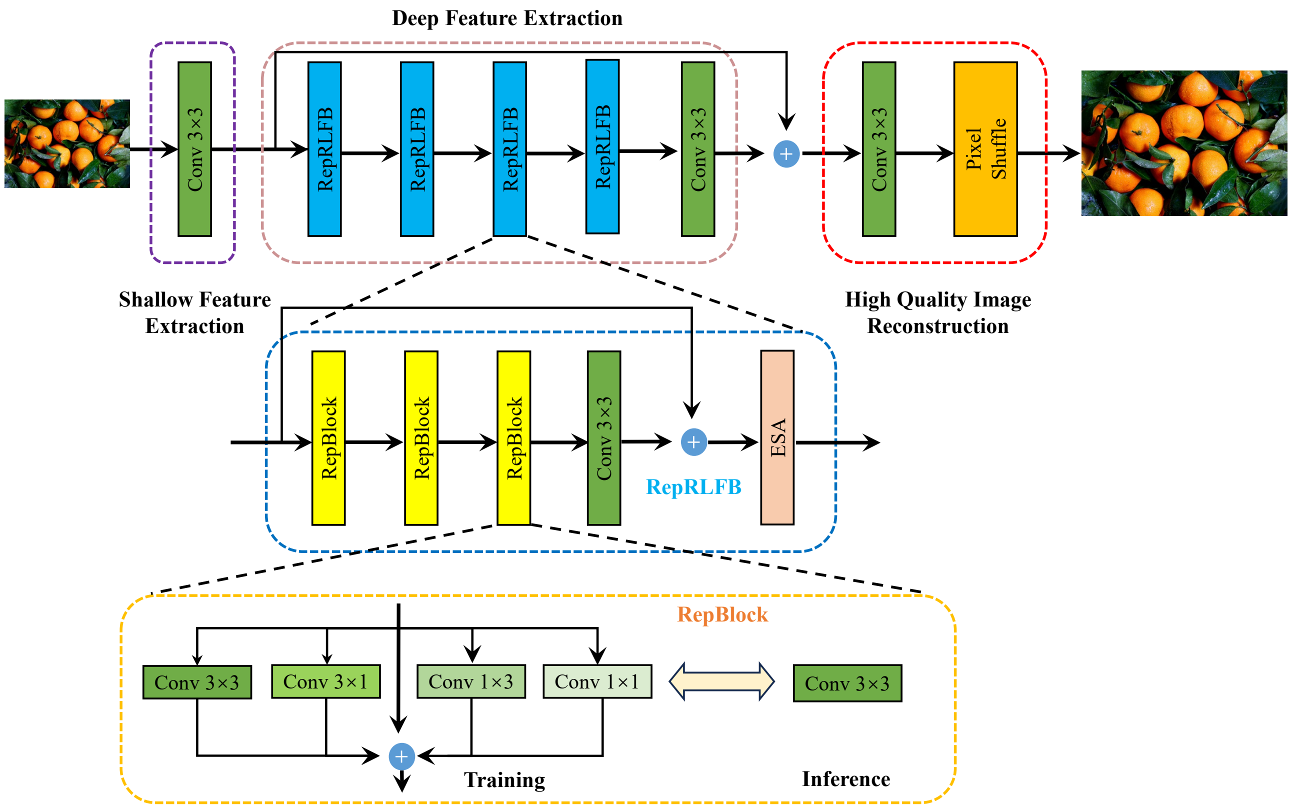

Figure 1. The network architecture of RepRLFN

transformer. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 22367–22377, 2023. 1

[2] Weijian Deng, Hongjie Yuan, Lunhui Deng, and Zengtong Lu. Reparameterized residual feature network for lightweight image super-resolution. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 1712–1721, 2023. 1

[3] Xiaohan Ding, Yuchen Guo, Guiguang Ding, and Jungong Han. Acnet: Strengthening the kernel skeletons for powerful cnn via asymmetric convolution blocks. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 1911–1920, 2019. 1

[4] Fangyuan Kong, Mingxi Li, Songwei Liu, Ding Liu, Jingwen He, Yang Bai, Fangmin Chen, and Lean Fu. Residual local feature network for efficient super-resolution. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 766–776, 2022. 1