Residual Fractal Network for Single Image Super Resolution by Widening and Deepening

Jiahang Gu, Zhaowei Qu, Xiaoru Wang, Jiawang Dan and Junwei Sun
Beijing Key Laboratory of Network System and Network Culture, Beijing University of Posts and Telecommunications
{gujiahang, zwqu, wxr, flydan, junweisun}@bupt.edu.cn

Problem
The repetitive texture features in the image play an important role in the super-resolution reconstruction process, which are usually ignored by most recent method. In response to this problem, we propose a residual fractal structure to increase the representation ability of the network on key internal features.

Basic Module
We design the residual fractal structure as a multi-branch convolution module. The difference of receptive field between adjacent branches is doubled. In the inter-branch fusion stage, 1x1 convolution is used to extract the features that appear in different branches as key features and participate in the final reconstruction stage. The difference with MSRN and RDN’s multi-scale structure is shown in the figure below.

Network Architecture
WRFN We use the recursive characteristics of RFCB to increase the number of branches to 7 to get WRFN, and reduce the difficulty of training through local and global skip connection. The structure with more branches can extract and integrate more levels of features, which is also a way of broadening the network.

DRFN In order to compare the effectiveness of the multi-scale structure, we also designed DRFN with reference to the network structure of MSRN and RDN. DRFN stacks multiple RFCBs, and merges the convolution outputs of different depths together to participate in reconstruction through the gate unit.

Quantitative Results by PSNR/SSIM
The comparison results for \( \times 3 \) with BI(left) and BD(right) on five standard benchmark datasets are shown in the figure below. Our RFNs+ outperform other compared methods on all the datasets. Without self-ensemble, RFNs obtain better performance in most datasets. The results of experiments show that multi-branch structure in RFCB can effectively extract recurring texture feature and utilize the autocorrelation of images. In addition, the width and depth of the network are both important factors to improve performance.

References

Acknowledgements
This research is supported by the National Natural Science Foundation of China under Grant 61672408 and Grant 61976025.

Visual Quality and Model Size Analyses
In order to compare the capabilities of extracting multi-scale features fairly, we only compare the results with MSRN and RDN. The visual comparisons are shown in the figure below. It can be seen that DRFN can restore sharper texture in the image while others suffer from blurring artifacts.

The model size and performance are shown in figure right. Compared with RCAN and SAN, our RFN has fewer parameters, while achieves better performance, especially the amount of parameters of RFN is reduced by 30%. Compared with RDN and EDSR, the amount of RFNs’ parameters is significantly reduced while the performance is greatly improved, which means RFNs achieve a better trade-off between the amount of parameters and performance. The better performance of WRFN and DRFN further verifies the effectiveness of RFCB and deepening and widening the network are both effective ways to improve performance.

Comparison of model size and performance. Results are evaluated on Set5 (2×).