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

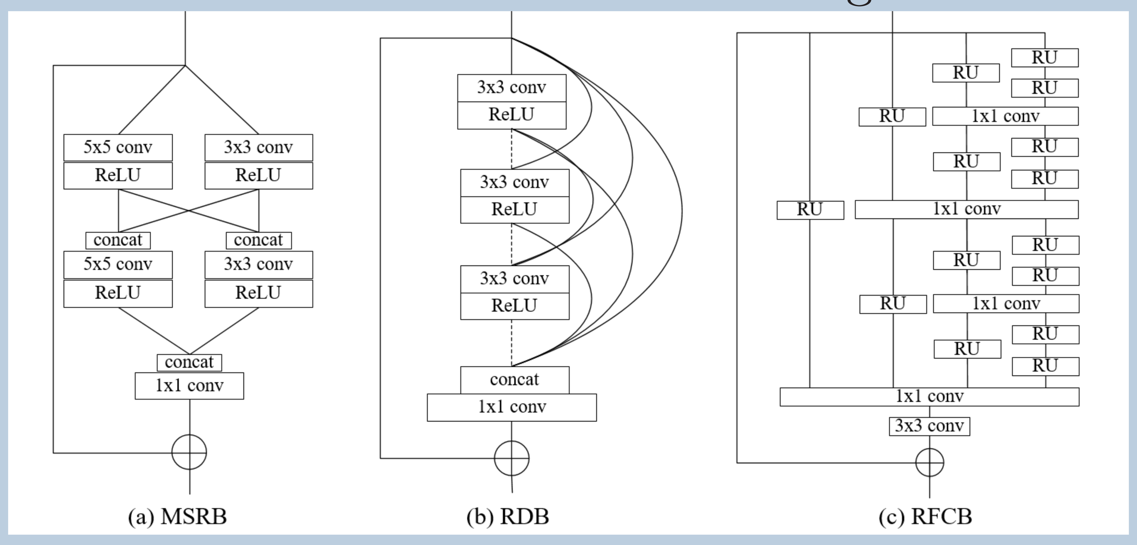
Jiahang Gu, Zhaowei Qu, Xiaoru Wang, Jiawang Dan and Junwei Sun
Beijing Key Laboratory of Network System and Network Culture, Beijing University of Posts and Telecommunications
{gujiahang, zwqu, wxr, flydan, junweisun}@bupt.edu.cn

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

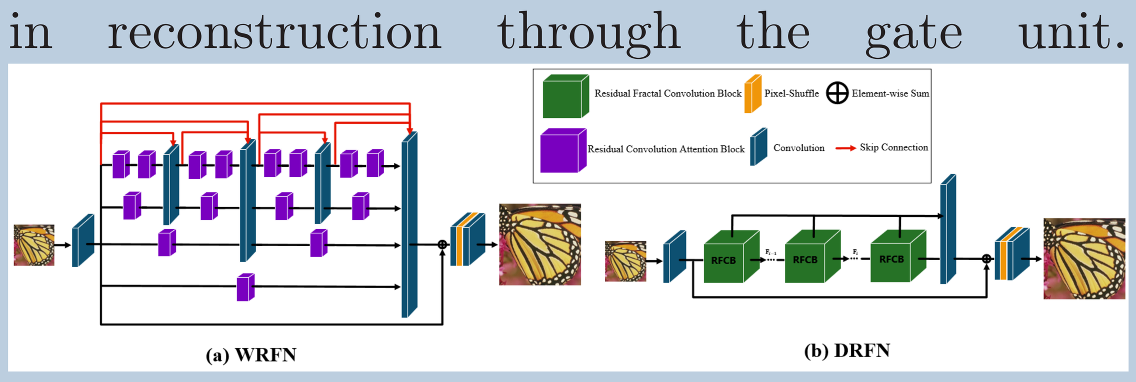
We design the residual fractal structure as a multi-branch convolution module. The difference of receptive field between adjacent branches is doubled. In the inter-branch fusion stage, 1x1 convolution is used to extract the features that appear in different branches as key features and participate in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



References

[1] Y. Zhang, Y. Tian, Y. Kong, B. Zhong, and Y. Fu *Residual dense network for image super-resolution*, in Proceedings of the IEEE conference on computer vision and pattern recognition, 2018, pp. 2472–2481.

[2] J. Li, F. Fang, K. Mei, and G. Zhang *Multi-scale residual network for image super-resolution*. in Proceedings of the European Conference on Computer Vision (ECCV), 2018, pp. 517–532.

[3] G. Larsson, M. Maire, and G. Shakhnarovich *Fractalnet: Ultra-deep neural networks without residuals*. arXiv preprint arXiv:1605.07648,2016.

Acknowledgements

This research is supported by the National Natural Science Foundation of China under Grant 61672108 and Grant 61976025.

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.

Method	Scale	Set5 PSNR/SSIM	Set14 PSNR/SSIM	B100 PSNR/SSIM	Urban100 PSNR/SSIM	Manga109 PSNR/SSIM
Bicubic	3	30.39/0.8682	27.55/0.7742	27.21/0.7385	24.46/0.7349	26.95/0.8556
SRCNN	3	32.75/0.9090	29.30/0.8215	28.41/0.7863	26.24/0.7989	30.48/0.9117
FSRCNN	3	33.18/0.9140	29.37/0.8240	28.53/0.7910	26.43/0.8080	31.10/0.9210
VDSR	3	33.67/0.9210	29.78/0.8320	28.83/0.7990	27.14/0.8290	32.01/0.9340
LapSRN	3	33.82/0.9227	29.87/0.8320	28.82/0.7980	27.07/0.8280	32.21/0.9350
MemNet	3	34.09/0.9248	30.01/0.8350	28.96/0.8001	27.56/0.8376	32.51/0.9369
EDSR	3	34.65/0.9280	30.52/0.8462	29.25/0.8093	28.80/0.8653	34.17/0.9476
MSRN	3	34.38/0.9262	30.34/0.8395	29.08/0.8041	28.08/0.8554	33.44/0.9427
RDN	3	34.71/0.9296	30.57/0.8468	29.26/0.8093	28.80/0.8653	34.13/0.9484
RCAN	3	34.74/0.9299	30.64/0.8481	29.32/0.8111	29.08/0.8702	34.43/0.9498
SAN	3	34.75/0.9300	30.59/0.8476	29.33/0.8112	28.93/0.8671	34.30/0.9494
DRFN	3	34.73/0.9301	30.64/0.8482	29.32/0.8108	28.99/0.8690	34.45/0.9497
DRFN+	3	**34.85/0.9307**	**30.75/0.8495**	29.36/0.8117	29.19/**0.8719**	**34.72/0.9510**
WRFN	3	34.77/0.9303	30.66/0.8483	29.33/0.8110	29.06/0.8694	34.45/0.9498
WRFN+	3	**34.85/0.9307**	30.73/0.8491	**29.37/0.8118**	**29.22**/0.8715	34.67/0.9509

Results With BI Degradation Model

Method	Scale	Set5 PSNR/SSIM	Set14 PSNR/SSIM	B100 PSNR/SSIM	Urban100 PSNR/SSIM	Manga109 PSNR/SSIM
Bicubic	3	28.78/0.8308	26.38/0.7271	26.33/0.6918	23.52/0.6862	25.46/0.8149
SPMSR	3	32.21/0.9001	28.89/0.8105	28.13/0.7740	25.84/0.7856	29.64/0.9003
SRCNN	3	32.05/0.8944	28.80/0.8074	28.13/0.7736	25.70/0.7770	29.47/0.8924
FSRCNN	3	26.23/0.8124	24.44/0.7106	24.86/0.6832	22.04/0.6745	23.04/0.7927
VDSR	3	33.25/0.9150	29.46/0.8244	28.57/0.7893	26.61/0.8136	31.06/0.9234
IRCNN	3	33.38/0.9182	29.63/0.8281	28.65/0.7922	26.77/0.8154	31.15/0.9245
SRMD	3	34.01/0.9242	30.11/0.8364	28.98/0.8009	27.50/0.8370	32.97/0.9391
RDN	3	34.58/0.9280	30.53/0.8447	29.23/0.8079	28.46/0.8582	33.97/0.9465
RCAN	3	34.70/0.9288	30.63/0.8462	29.32/0.8093	28.81/0.8645	34.38/0.9483
SAN	3	34.75/0.9290	30.68/0.8466	29.33/0.8101	28.83/0.8646	34.46/0.9487
DRFN	3	34.73/0.9295	30.68/0.8487	29.34/0.8108	28.91/0.8665	34.58/0.9494
DRFN+	3	**34.87/0.9302**	**30.79/0.8487**	**29.39/0.8117**	29.10/**0.8691**	**34.85**/0.9506
WRFN	3	34.77/0.9297	30.69/0.8473	29.35/0.8106	28.97/0.8665	34.63/0.9496
WRFN+	3	34.84/0.9301	30.78/0.8484	29.39/0.8113	**29.13**/0.8688	34.84/**0.9507**

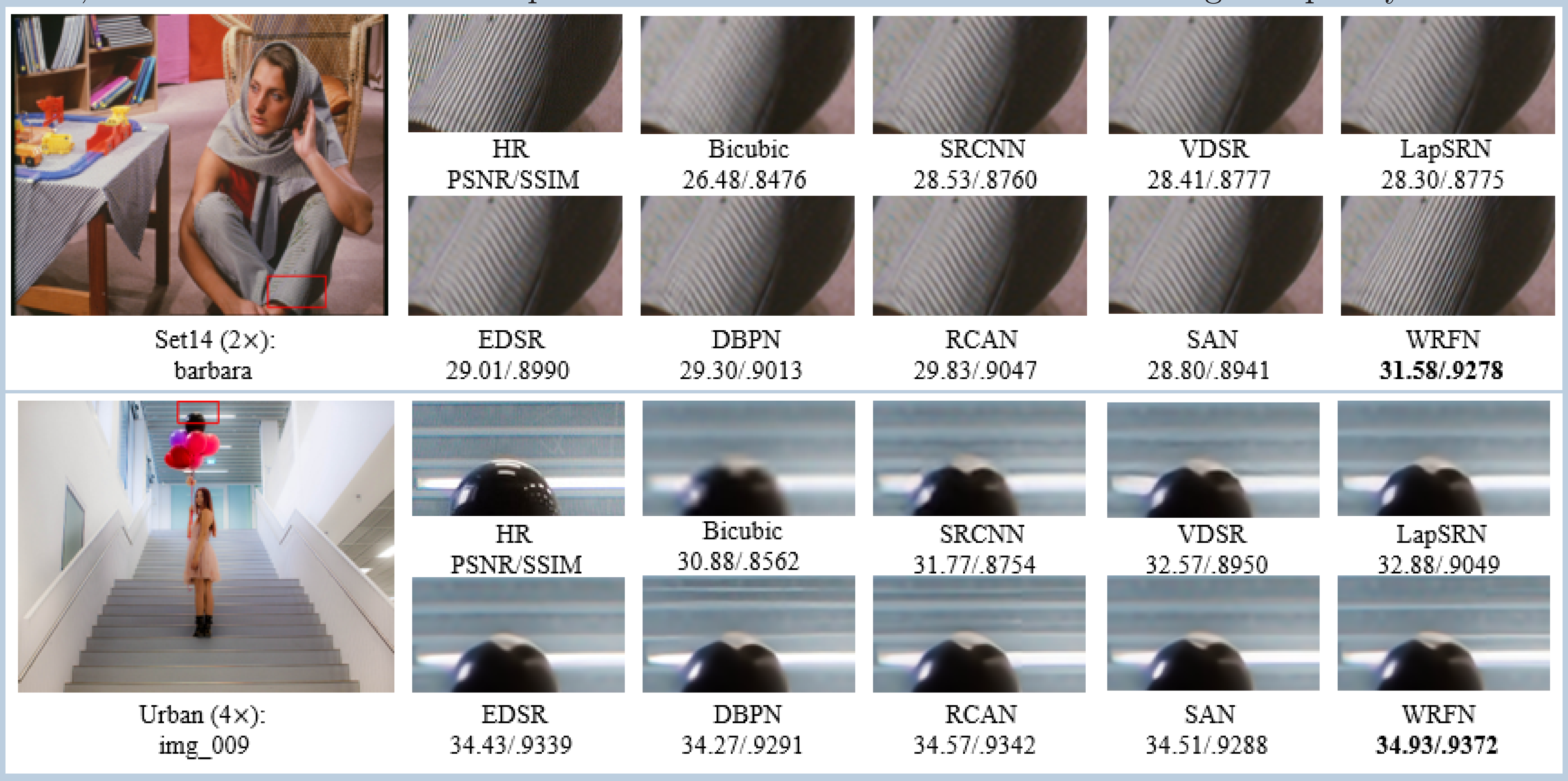
Results With BD Degradation Model

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.

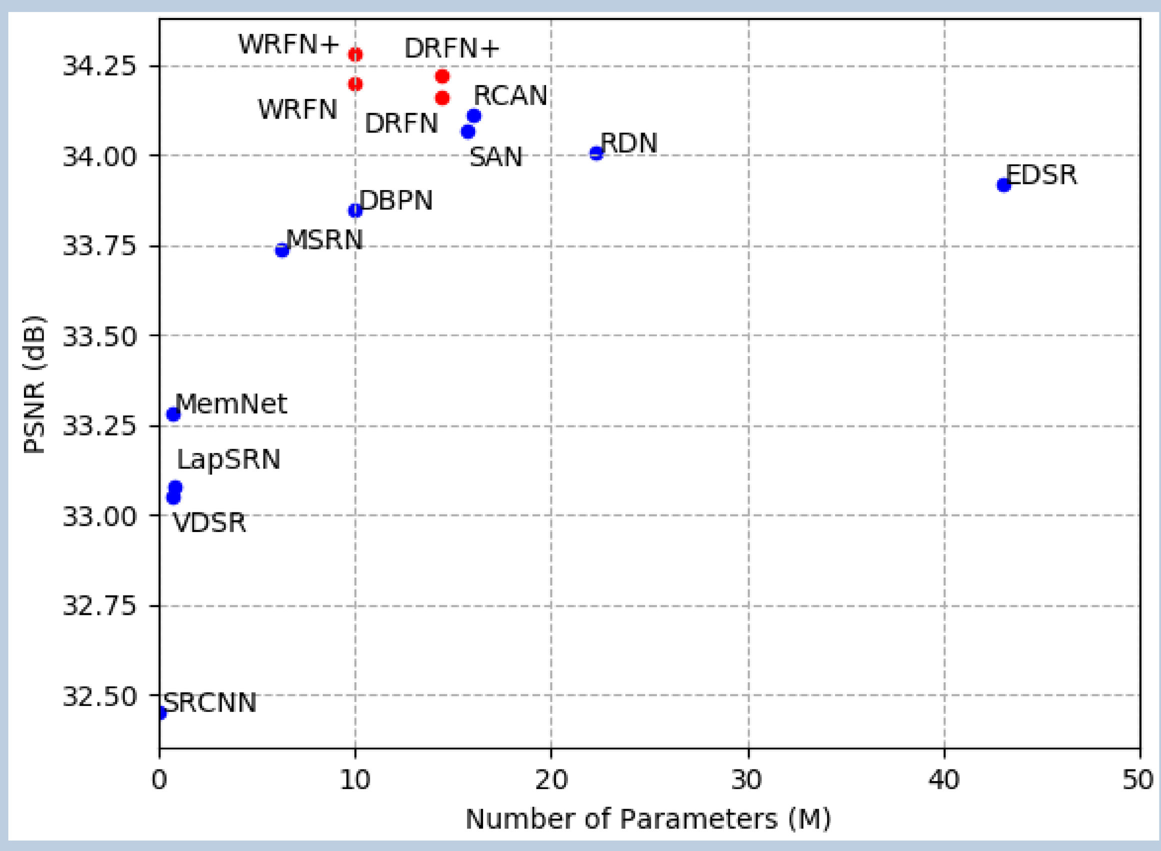
Visual comparison for $\times 2$ SR with BI model on Set14 dataset.

The visual comparisons of WRFN and previous methods are shown in the following figures, from which we can see most compared models cannot reconstruct texture accurately and suffer from serious blurring artifacts. In contrast, WRFN obtains sharper texture and restores more high-frequency details.



Visual comparison with BI model on Set14 ($2\times$) and Urban100 ($4\times$).

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 WRFN 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 ($2\times$)