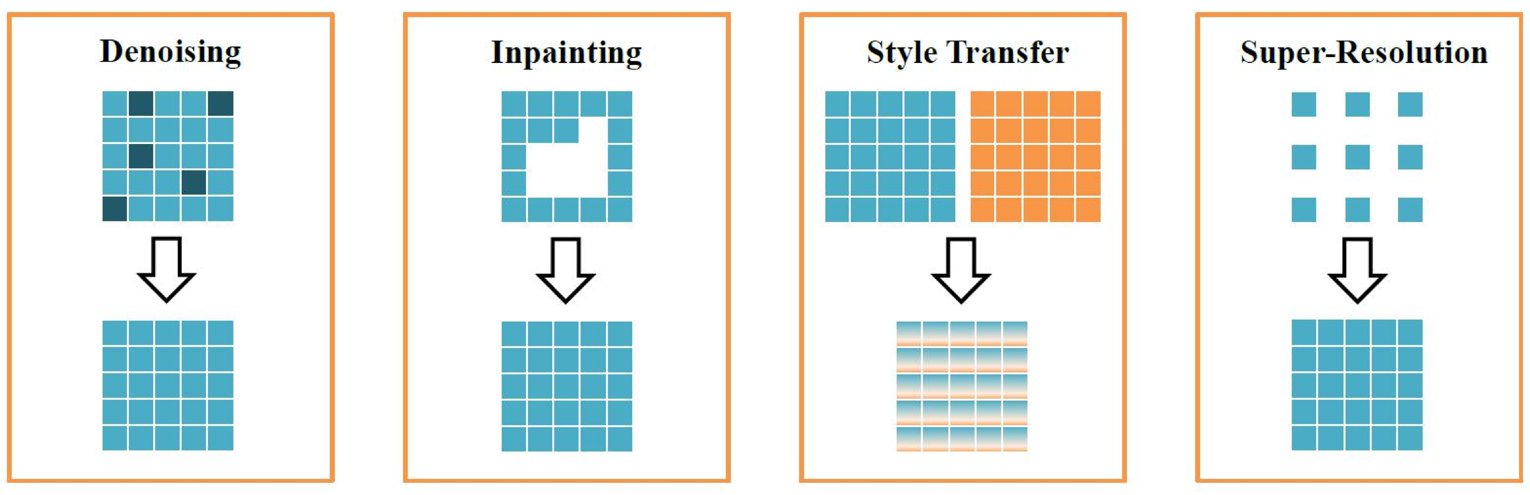
Cross-layer Information Refining Network for Single Image Super-Resolution

Hongyi Zhang, Wen Lu, Xiaopeng Sun

Xidian University, Xi'an, China

Introduction

Single Image Super-Resolution



Contributions

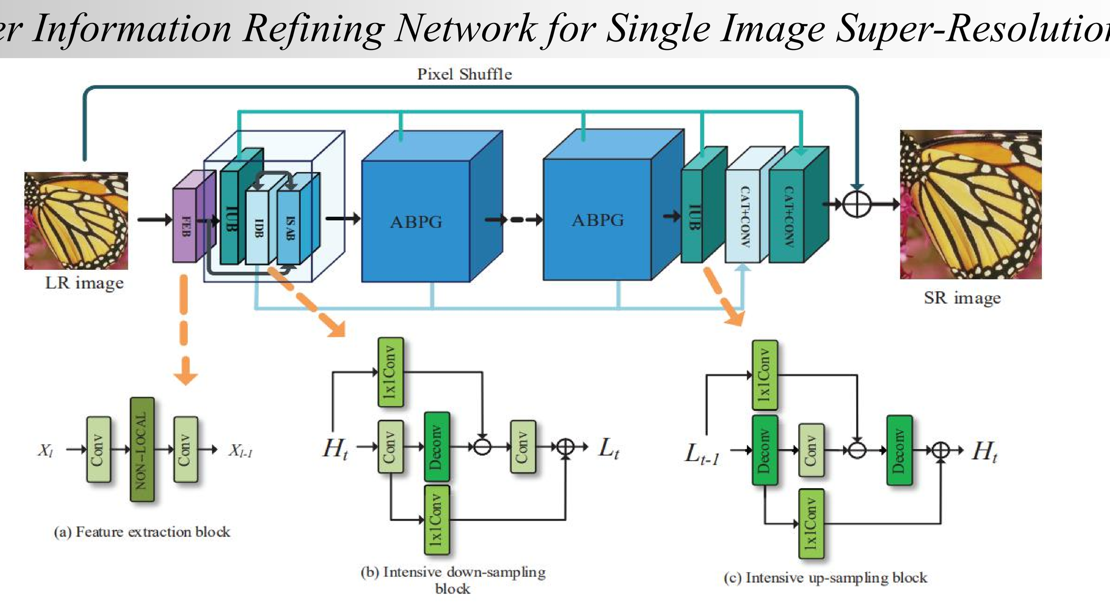
(1) Without adding too many parameters, a novel interlaced spatial attention block is proposed to adaptively refine features by considering correlations among adjacent layers. The proposed ISAB module incorporates manifold information, which enhances the diversity of network topology structure and representational power.

(2) By applying ISAB, we utilize attention based back projection block to build a deep network for image SR. Experimental results demonstrate that our method is substantially more accurate. Owing to information refinement in adjacent alayers, our network still perform well at large scale magnifications.

Our Method

General Framework

Cross-layer Information Refining Network for Single Image Super-Resolution

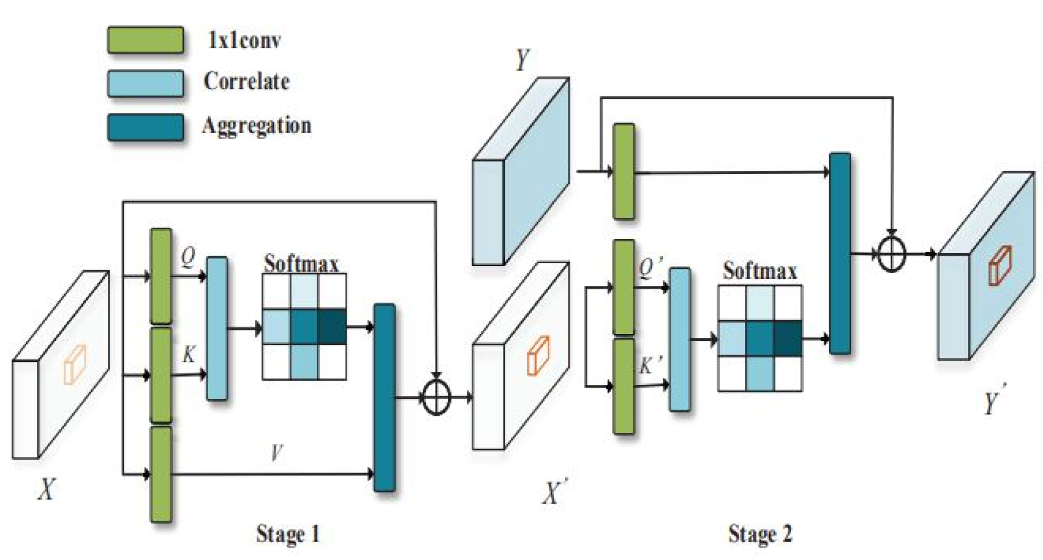


The proposed CIRN consists of feature extraction block(FEB), n attention-based back projection groups(ABPG). Every ABPG consists of intensive down-sampling block(IDB), intensive up-sampling block(IUB) and interlaced spatial attention block(ISAB).

Details on Interlaced Spatial Attention Block

● *Dilated convolution & Multi-order gradients*

Inspired by the non-local module and the CCNet, ISAB adopts the idea of divide and conquer, which reduces the amount of calculation while capturing the relationship between the feature maps of adjacent layers.



Let's X denotes the feature map from previous layer, and Y denotes current layer's feature map. In this way, we have

$$X' = \sum_{I=1}^{H+W-1} softmax(K_m Q_m)V_m + X_m$$

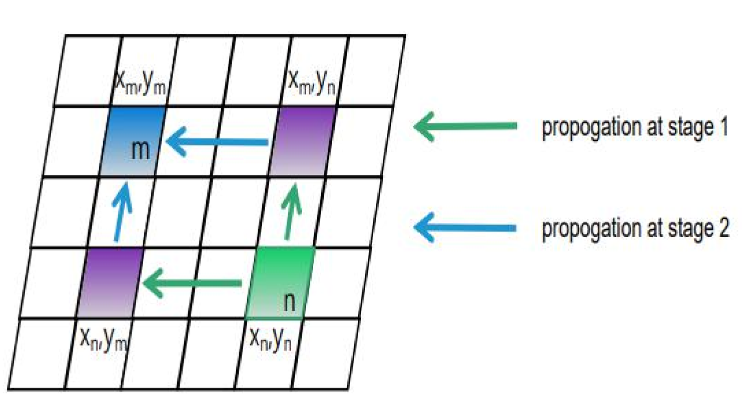
At the stage 1, the pixel in position m captures sparse self-attention map from one feature map and the second stage generate global attention map without adding extra parameters through interlace operation

● *Interlace operation*

The interlace operation can be formulated as :

$$Y_m = f_x(S_1, X_n, Y_n) + f_y(S_2, X_n, Y_m) + f_y(S_1, X_n, Y_n) + f_x(S_2, X_m, Y_n)$$

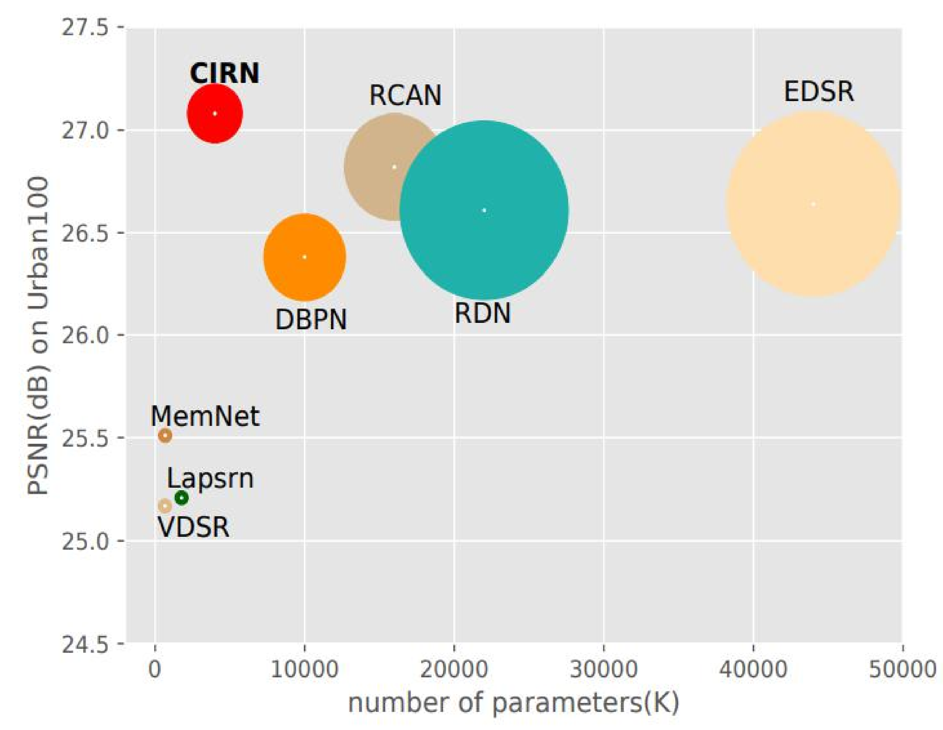
in which Y_m denotes the attention map in position m, $f_x(S_k, X_n, Y_n)$represents the function that position n information spread in x-axis direction at stage k.



Apparently we can see that any position, even not in vertical or horizontal of m, will finally propagate information to position m through interlace operation. These operations in two stages share the same parameters to avoid adding too many extra parameters. After obtained the attention map with rich contextual information, we utilize the feature maps from the current layer to aggregate with it. Owing to the interlaced spatial attention mechsiam, information from adjacent layers is fused to obtain more useful information for reconstruction.

Model Analysis

● *Performance and number of parameters*



We conducted a comparative experiment on the parameters and performance of the model on the Urban100(4 ×) dataset. The reference works are the methods which achieve good SR results in the past two years. The experimental results show that our proposed CIRN achieves high performance while maintaining a lighter model size.

● *Effect of ISAB*

We visualized the input and output feature maps of ISAB. Specifically X,Y represent the input feature maps while Y' represents the output feature maps. It is evident that output Y' refined the information of inputs and enhanced the edges information from X which is exactly Y lost.

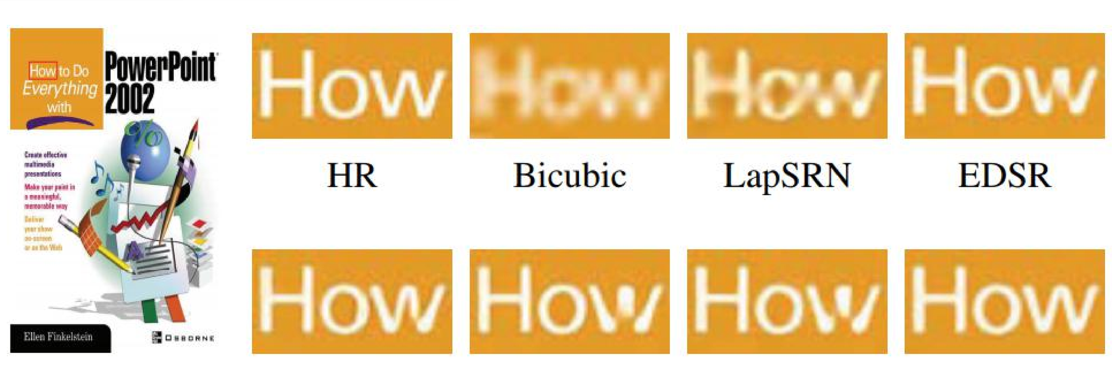
Experiment Results

Quantitative Results

Algorithm	Scale	Set5 PSNR	Set5 SSIM	Set14 PSNR	Set14 SSIM	BSD 100 PSNR	BSD 100 SSIM	Urban 100 PSNR	Urban 100 SSIM	Manga109 PSNR	Manga109 SSIM
Bicubic	4x	28.42	0.810	26.10	0.704	25.96	0.669	23.64	0.659	25.15	0.789
SRCNN [1]		30.49	0.862	27.61	0.754	26.91	0.712	24.89	0.744	28.12	0.872
LapSRN [24]		31.54	0.855	28.19	0.772	27.32	0.728	25.21	0.756	29.09	0.890
EDSR [5]		32.46	0.897	28.80	0,788	27.71	0.742	26.64	0.803	31.02	0.915
RDN [8]		32.47	0.899	28.81	0.787	27.72	0.741	26.61	0.802	31.00	0.915
DBPN [15]		32.47	0.898	28.82	0.786	27.72	0.740	26.38	0.794	30.91	0.913
RCAN [13]		32.63	0.900	28.87	0.788	27.77	0.743	26.82	0.808	31.22	0.917
SAN [14]		32.64	0.900	28.92	0.789	27.78	0.743	26.79	0.806	31.18	0.916
CIRN(Ours)		32.70	0.901	28.98	0.790	27.82	0.744	27.08	0.811	31.72	0.920
Bicubic	8x	24.40	0.658	23.10	0.566	23.67	0.548	20.74	0.516	21.47	0.650
SRCNN [1]		25.33	0.690	23.76	0.591	24.13	0.566	21.29	0.544	22.46	0.695
LapSRN [24]		26.15	0.738	24.35	0.620	24.54	0.586	21.81	0.581	23.39	0.735
EDSR [5]		26.96	0.776	24.91	0.642	24.81	0.598	22.51	0.622	24.69	0.784
HBPN [25]		27.17	0.785	24.96	0.642	24.93	0.602	23.04	0.647	25.24	0.802
DBPN [15]		27.21	0.784	25.13	0.648	24.88	0.601	22.73	0.631	25.14	0.798
RCAN [13]		27.31	0.787	25.23	0.651	24.98	0.605	23.00	0.645	25.24	0.803
SAN [14]		27.22	0.782	25.14	0.647	24.88	0.601	22.70	0.631	24.85	0.790
CIRN(Ours)		27.37	0.788	25.32	0.652	25.00	0.607	23.10	0.646	25.47	0.808
Bicubic	16x	-	-	-	-	21.71	0.477	18.92	0.434	19.10	0.568
EDSR [5]		-	-	-	-	22.62	0.506	19.96	0.481	20.62	0.635
RCAN [13]		-	-	-	-	22.69	0.511	20.20	0.496	20.88	0.656
SAN [14]		-	-	-	-	22.86	0.519	20.27	0.508	21.06	0.666
CIRN(Ours)		-	-	-	-	22.90	0.521	20.43	0.517	21.25	0.675

In objective experiments, it can be seen that our method is superior to the previous method in both PSNR and SSIM. In addition, our method still performs well in the case of extreme magnification.

Qualitative Results



HR Bicubic LapSRN EDSR

GroundTruth RDN RCAN SAN CIRN(ours)

"*ppt3*"(4x)

HR Bicubic EDSR

GroundTruth RCAN SAN CIRN(ours)

"*Image003*"(16x)

Acknowledgements

This research was supported in part by the National Natural Science Foundation of China (Grant Nos.61871311, 61876146),the Key Industrial Innovation Chain Project in Industrial Domain of Shaanxi Province (Grant No. 2020ZDLGY05-01), the National Key Research and Development Program of China (No.2018AAA0102702).

Contact Information

Hongyi Zhang

Email: 542490069@qq.com,
Hyzhang_12@stu.xidian.edu.cn