ICIP2021 Anchorage, Alaska.

浙江師範大學 ZHEJIANG NORMAL UNIVERSITY

Single Image Super-resolution via Global-context Attention Networks


Pengcheng Bian


Zhonlong Zheng


Dawei Zhang


Liyuan Chen


Minglu Li

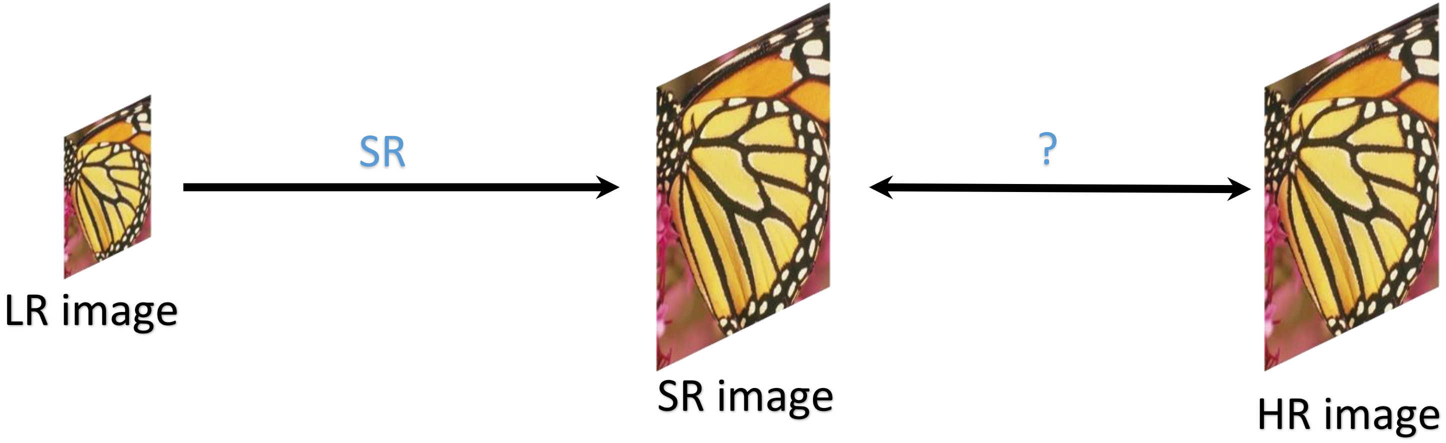
Presenter: Pengcheng Bian

College of Mathematics and Computer Science, Zhejiang Normal University, China

Introduction: Single Image Super-resolution

- **Single Image Super-resolution (SISR):** reconstruct a high-resolution (HR) image from the corresponding low-resolution (LR) image.

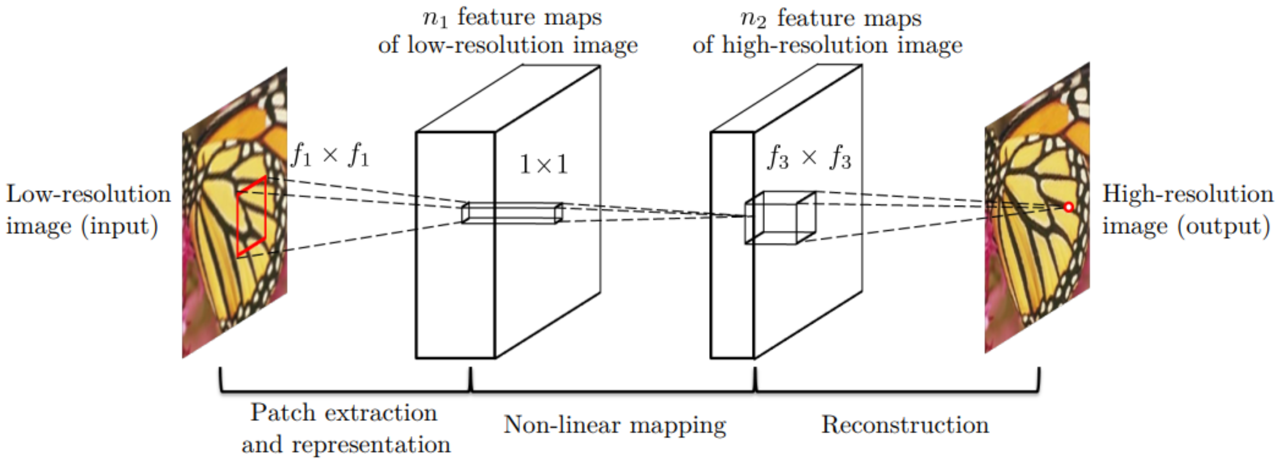
- Given a LR image with the size of $W \times H \times \mathrm{C}$, obtain a HR image with the size of $sW \times sH \times C$ via super-resolution, s is the scale factor.



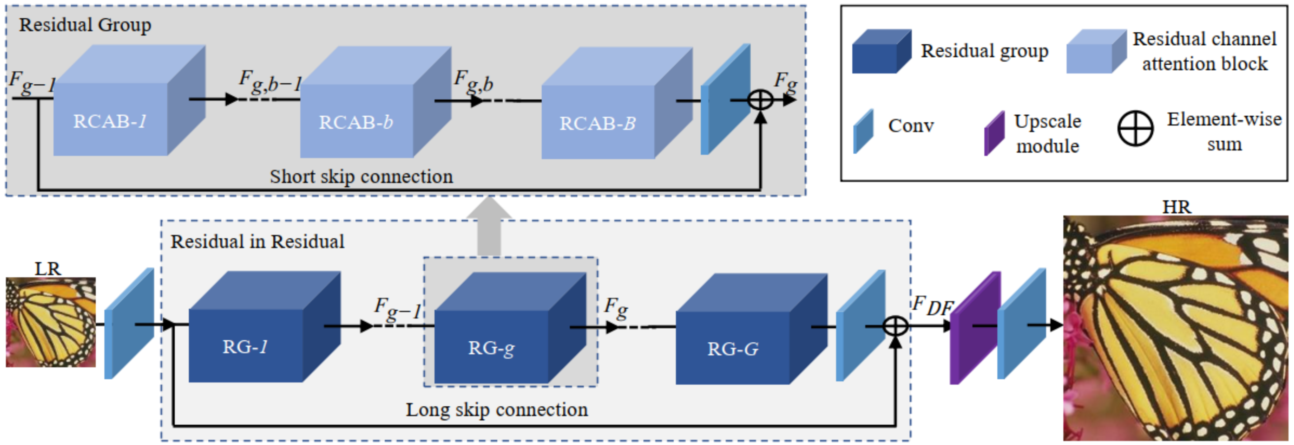
Introduction: Single Image Super-resolution

- **SISR based on Convolutional Neural Networks:**

SRCNN[1]



RCAN[2]



[1] C. Dong, C. C. Loy, K. He, and X. Tang, "Learning a deep convolutional network for image super-resolution," in *ECCV*, 2014.
[2] Y. Zhang, Li K, K. Li, L. Wang, B. Zhong, and Y. Fu, "Image super-resolution using very deep residual channel attention networks," in *ECCV*, 2018.

Motivations

1. As the network depth grows, the features in each layer contribute differently to the final reconstruction.

2. Most existing CNN-based methods use one or more types of attention independently, resulting in insufficient modeling for multi-dimensional correlations.

Proposed Method

Architecture of Global-context Attention Networks

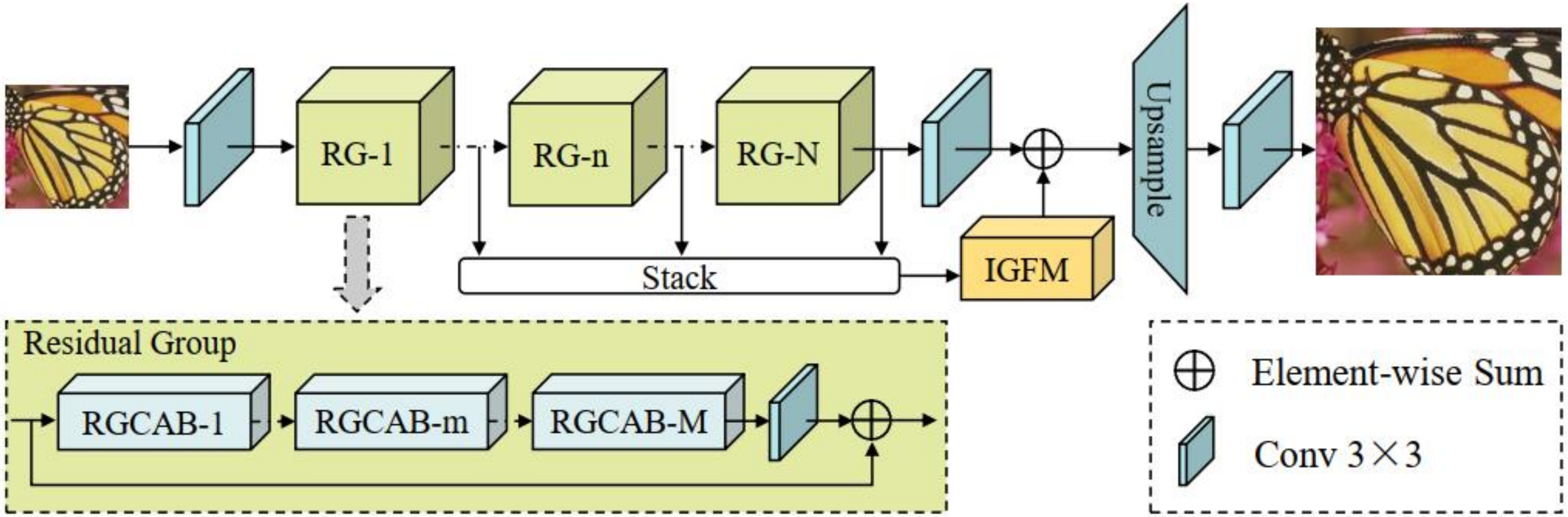


Fig. 1. Network architecture of the proposed global-context attention network (GCAN). The network consists of a set of residual groups, each of which contains several residual global-context attention blocks (RGCABs), and the outputs of all residual groups are stacked together as the input of the inter-group fusion module (IGFM) to explore correlations among group features. Finally, the resulting features are up-sampled to get the final high-resolution image.

Proposed Method

Sub-modules

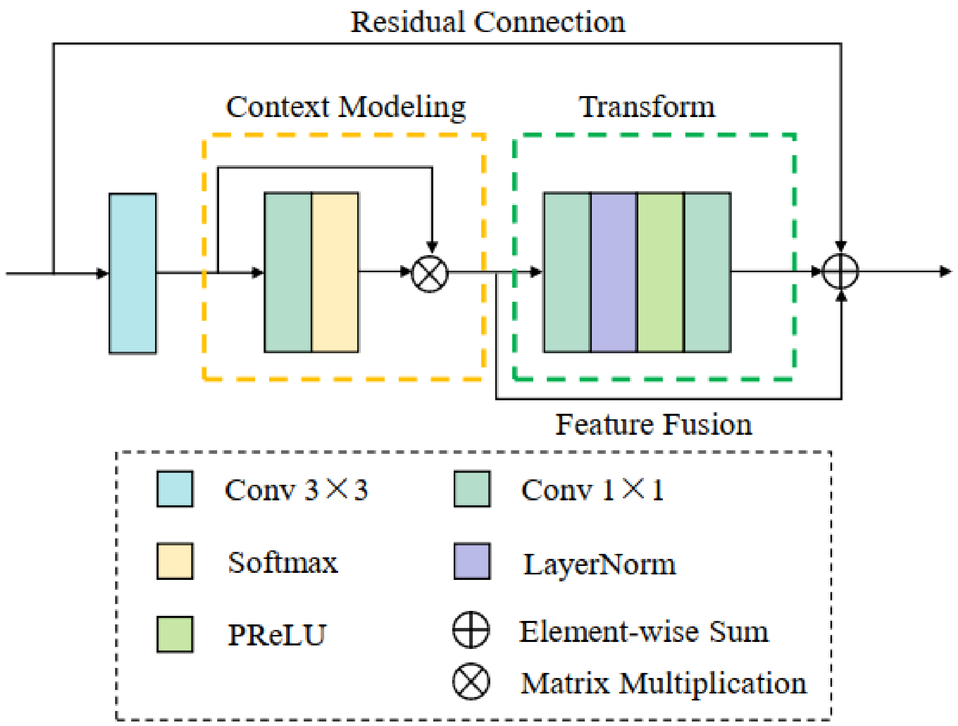


Fig. 2. Residual Global-Context Attention Block (RGCAB).

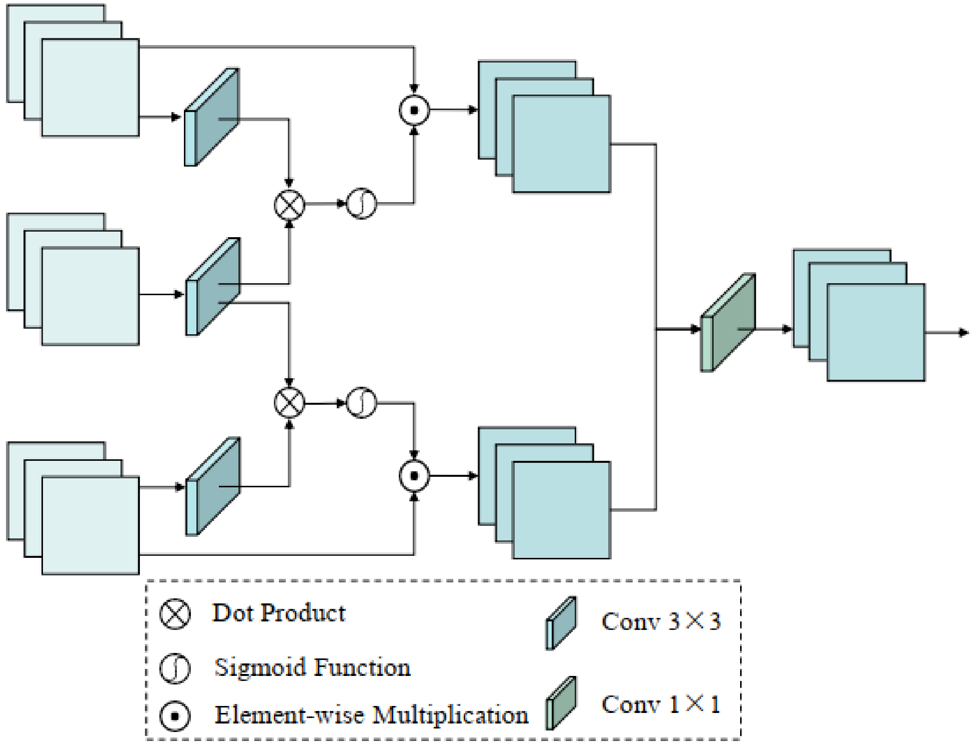


Fig. 3. Inter-Group Fusion Module (IGFM).

Network Training

Datasets- Training set: DIV2K

- Testing set: Set5, Set14, B100, Urban100

Loss Function: $L(\Theta) = \frac{1}{k}\sum_{i=1}^{k}\|H_{GCAN}(I^i_{LR}) - I^i_{HR}\|_1$

Evaluation Metrics: PSNR and SSIM

Experimental Results

Table 1. Quantitative results on four datasets: average PSNR/SSIM results for scale factors $2\times$, $3\times$ and $4\times$.

Methods	Scale	Set5		Set14		B100		Urban100	
		PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
Bicubic	×2	33.66	0.9299	30.24	0.8688	29.56	0.8431	26.88	0.8403
SRCNN	×2	36.66	0.9542	32.45	0.9067	31.36	0.8879	29.50	0.8946
LapSRN	×2	37.52	0.9591	33.08	0.9130	31.08	0.8950	30.41	0.9101
EDSR	×2	38.11	0.9602	33.92	0.9195	32.32	0.9013	32.93	0.9351
DBPN	×2	38.09	0.9600	33.85	0.9190	32.27	0.9000	32.55	0.9324
RCAN	×2	38.27	0.9614	34.12	0.9216	32.41	0.9027	33.34	0.9384
SRFBN	×2	38.11	0.9609	33.82	0.9196	33.29	0.9010	32.62	0.9328
GCAN	×2	38.28	0.9615	34.15	0.9217	32.43	0.9029	33.38	0.9386
GCAN+	×2	**38.32**	**0.9617**	**34.22**	**0.9221**	**32.46**	**0.9031**	**33.49**	**0.9392**
Bicubic	×3	30.39	0.8682	27.55	0.7742	27.21	0.7385	24.46	0.7349
SRCNN	×3	32.75	0.9090	29.30	0.8215	28.41	0.7863	26.24	0.7989
LapSRN	×3	33.82	0.9227	29.87	0.8320	28.82	0.7980	27.07	0.8280
EDSR	×3	34.65	0.9280	30.52	0.8462	29.25	0.8093	28.80	0.8653
RCAN	×3	34.74	0.9299	30.65	0.8482	29.32	0.8111	29.09	0.8702
SRFBN	×3	34.70	0.9292	30.51	0.8461	29.24	0.8084	28.73	0.8641
GCAN	×3	34.79	0.9300	30.67	0.8483	29.34	0.8113	29.10	0.8703
GCAN+	×3	**34.85**	**0.9303**	**30.76**	**0.8493**	**29.40**	**0.8120**	**29.25**	**0.8722**
Bicubic	×4	28.42	0.8104	26.00	0.7027	25.96	0.6675	23.14	0.6577
SRCNN	×4	30.48	0.8628	27.50	0.7513	26.90	0.7101	24.52	0.7221
LapSRN	×4	31.54	0.8850	28.19	0.7720	27.32	0.7270	25.21	0.7560
EDSR	×4	32.46	0.8968	28.80	0.7876	27.71	0.7420	26.64	0.8033
DBPN	×4	32.47	0.8980	28.82	0.7860	27.72	0.7400	26.38	0.7946
RCAN	×4	32.63	0.9002	28.87	0.7889	27.77	0.7436	26.82	0.8087
SRFBN	×4	32.47	0.8983	28.81	0.7868	27.72	0.7409	26.60	0.8015
GCAN	×4	32.64	0.9003	28.91	0.7889	27.79	0.7437	26.83	0.8088
GCAN+	×4	**32.72**	**0.9011**	**28.98**	**0.7901**	**27.84**	**0.7450**	**26.98**	**0.8126**

Visual comparisons

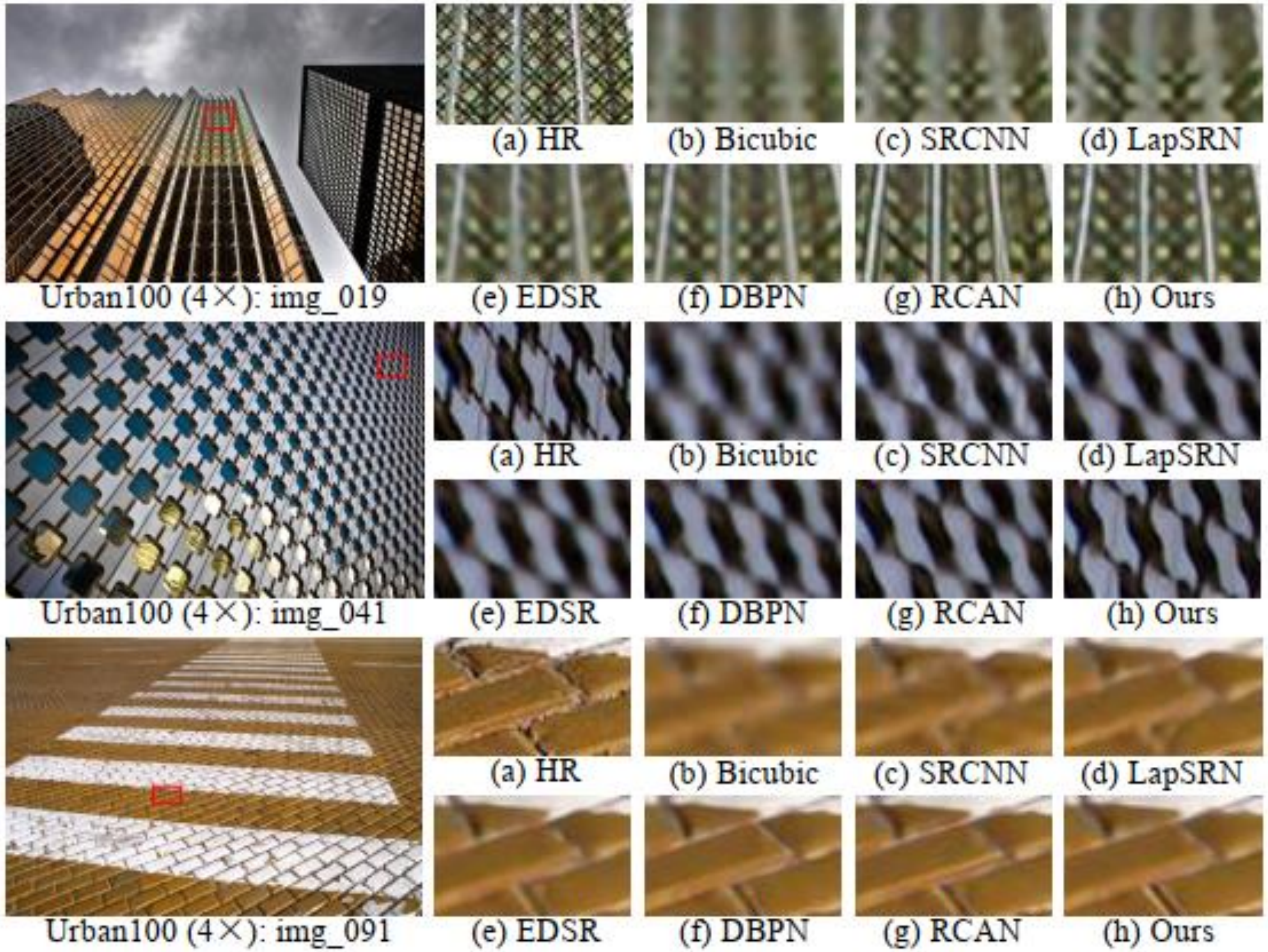


Fig. 4. Visual comparisons for $4\times$ SR on Urban100 dataset.

Conclusion

- A global-context attention network (GCAN) is proposed for single image super-resolution (SISR). Taking advantage of efficient global-context attention, the multiple stacked residual attention blocks are able to explore long-range spatial correlations and channel interdependencies effectively.
- An inter-group feature fusion module (IGFM) is adopted to make full use of hierarchical features among residual groups. By learning relations of inter-group features, multiple representative features can be extracted and aggregated sufficiently.
- Extensive experiments on several public datasets demonstrate that our method performs favorably against the other SISR approaches in terms of accuracy and visual results.

Thanks for your time !