IVCL

Image Super-Resolution Using Very Deep Residual Channel Attention Networks

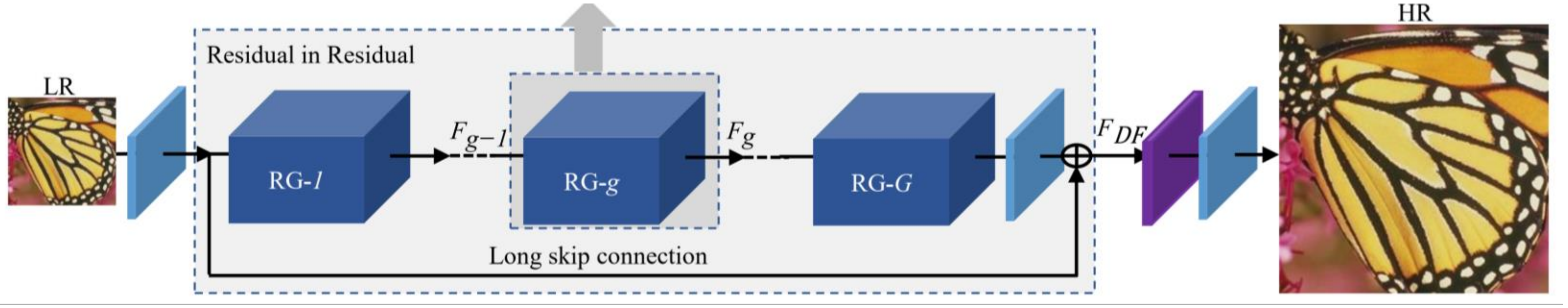
Yulun Zhang et al., 2018

Presenter: Le Van The

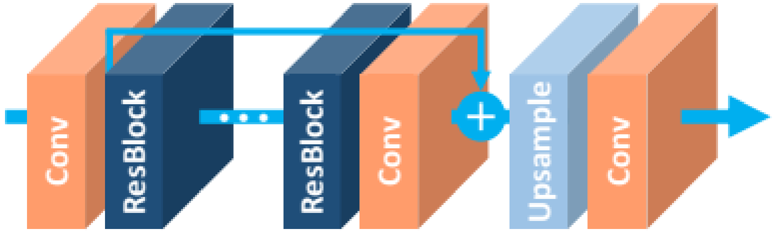
Outline

- Proposed RCAN
- Experiment results
- Conclusion

Proposed model

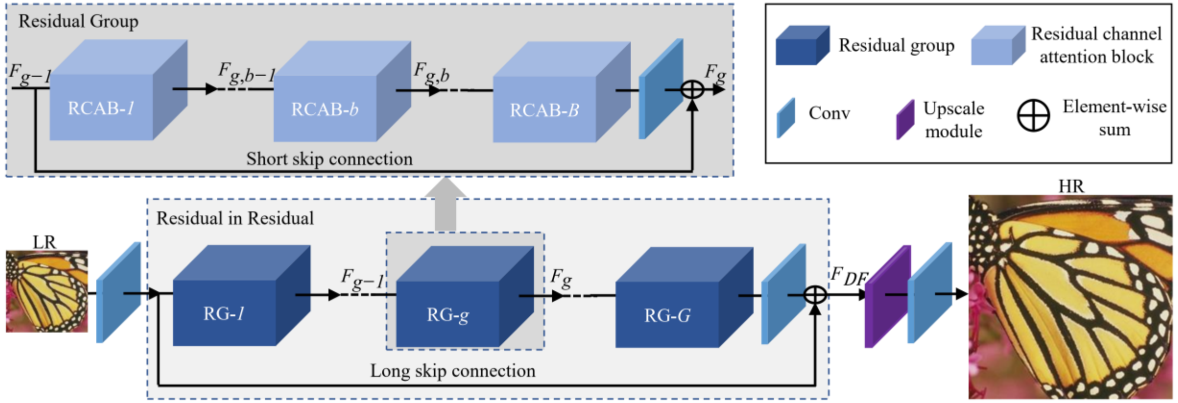


Residual channel attention network (RCAN)



Enhanced deep super-resolution (EDSR)

RCAN

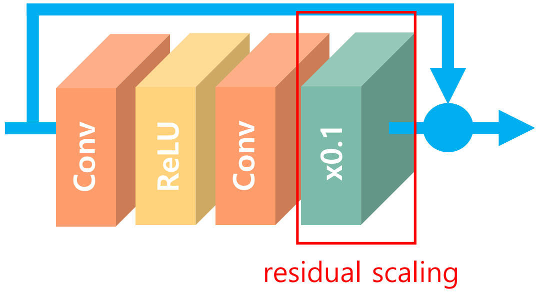


Residual channel attention network (RCAN)

❖ Residual in Residual (RIR):
  - G residual groups (RG) and long skip connection (LSC)
  - Each RG further contains B residual channel attention blocks (RCAB) with short skip connection (SSC)

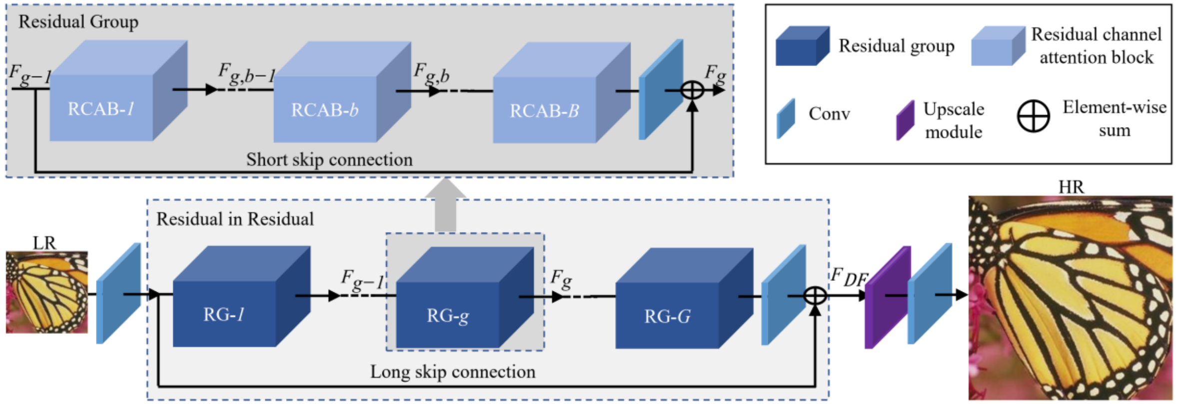
- Motivation :
  - Very deep feature extraction is hard for training
  - Long and short skip connection bypass abundant low-frequency info rmation → can ease the flow of information→ stabilize the training of very deep network

- Effect:
  - allows to train very deep CNN (over 400 layers) for image SR with hi gh performance
  - achieve more than 1,000 layers trainable networks



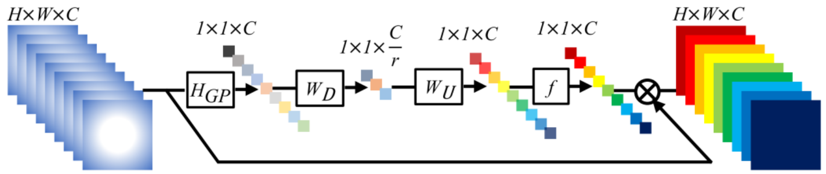
Residual scaling in EDSR

RCAN



Residual channel attention network (RCAN)

❖ Channel Attention (CA):



Channel attention (CA)

- Motivation :
  - Previous networks extract features from the original LR inputs and consider each channel-wise feature equally → wastes unnecessary computations for abundant low-frequency features

  → lack discriminative learning ability across feature channels

  → hinder the representational power of deep networks

- Effect:
  - adaptively rescale features

  →improves the representational ability of the network

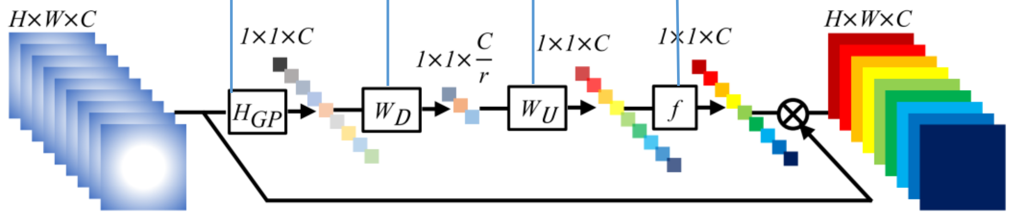
RCAN

❖ Channel Attention (CA):

Gating Mechanism

Conv layers + ReLu

Global Average Pooling



Channel attention (CA)

Global Average Pooling

$$z_c = H_{GP}(x_c) = \frac{1}{H \times W} \sum_{i=1}^{H} \sum_{j=1}^{W} x_c(i,j),$$

where x_c (i, j) is the value at position (i, j) of c-th feature x_c

1	3	4
4	3	1
2	2	2

x_c

$$z_c = \frac{1}{3\times3}(1+3+\cdots+2) = 2$$

Conv layers + ReLu

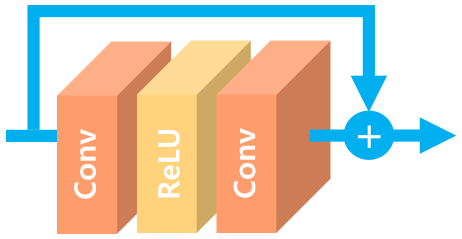
The bottleneck with two conv layers are formed with dimensionality reduction using reduction ratio *r*.

Gating Mechanism

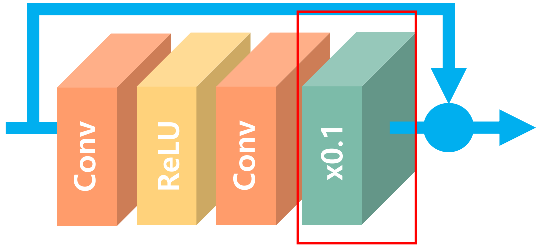
Use sigmoid function → capture channel-wise dependencies from the aggregated information

RCAB

❖ Residual channel attention block (RCAB)



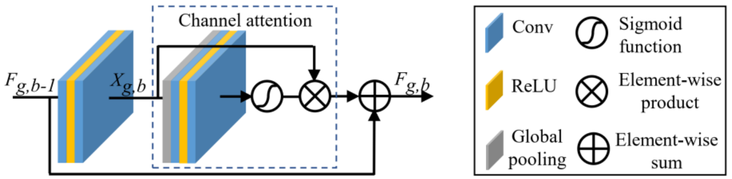
Residual block (RB) in EDSR



residual scaling

Residual scaling block (RCB) in EDSR

Integrate CA

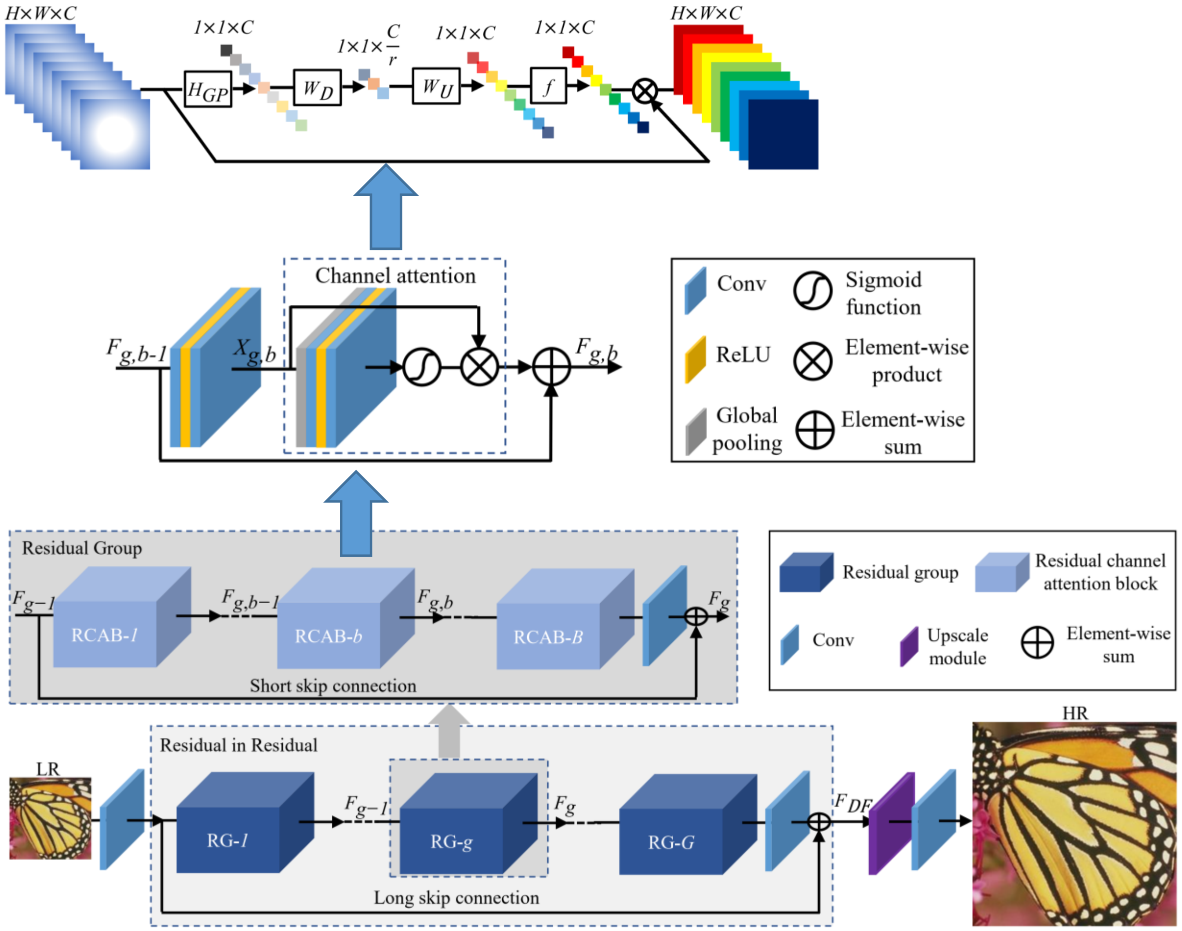


Residual channel attention block (RCAB)

RB and RCB are special cases of RCAB:

- When channel attention is 1 constant, RCAB=RB
- When channel attention is 0.1 constant, RCAB=RCB

RCAN



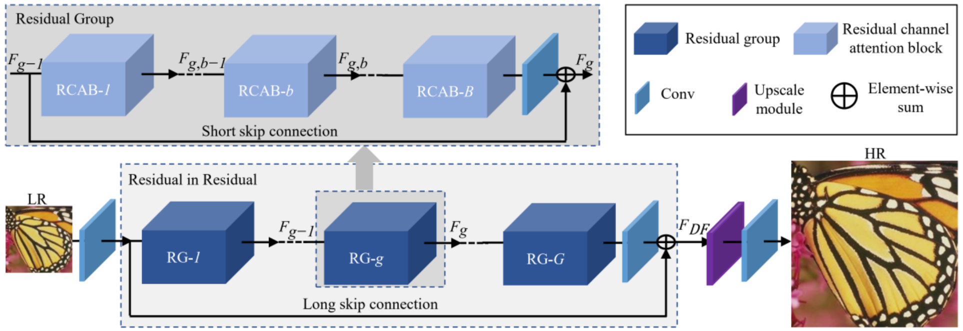
Channel attention
Conv
ReLU
Global pooling
Sigmoid function
Element-wise product
Element-wise sum
Residual Group
RCAB-1
RCAB-b
RCAB-B
Short skip connection
Residual group
Residual channel attention block
Upscale module
Residual in Residual
LR
RG-1
RG-g
RG-G
Long skip connection
HR

Experiment results

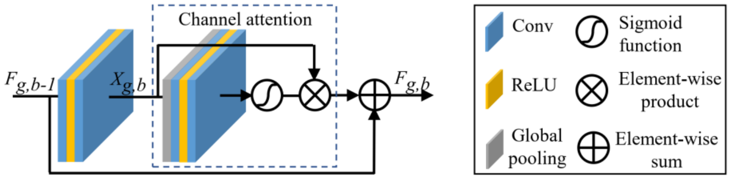
- Training setting:
  - Dataset: DIV2K (800 images)
  - Data augmentation: random flips and rotations
  - Learning rate: 1e-4, halved at every 2e5 iterations
  - Optimizer: ADAM
  - Loss: L1
- Model setting:
  - G=10, B=20
  - Reduction ratio r = 16

Table 1: Investigations of RIR including Long Skip Connection (LSC), Short Skip Connection (SSC), and Channel Attention (CA)

Residual in Residual (RIR)	LSC	✗	✓	✗	✓	✗	✓	✗	✓
	SSC	✗	✗	✓	✓	✗	✗	✓	✓
Channel attention (CA)		✗	✗	✗	✗	✓	✓	✓	✓
PSNR on Set5 (2×)		37.45	37.77	37.81	37.87	37.52	37.85	37.86	37.90

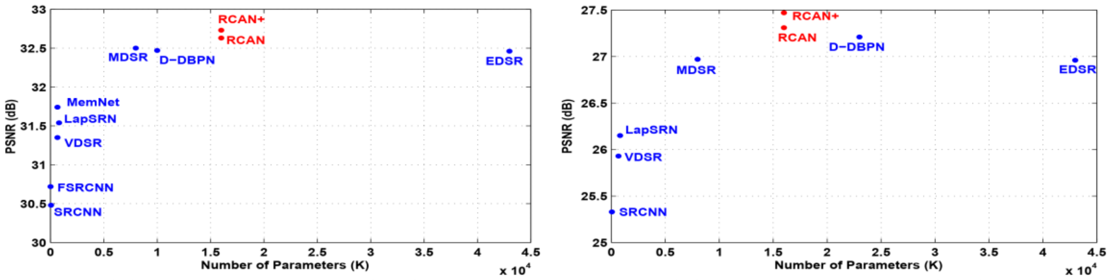


Residual channel attention network (RCAN)



Residual channel attention block (RCAB)

Experiment results

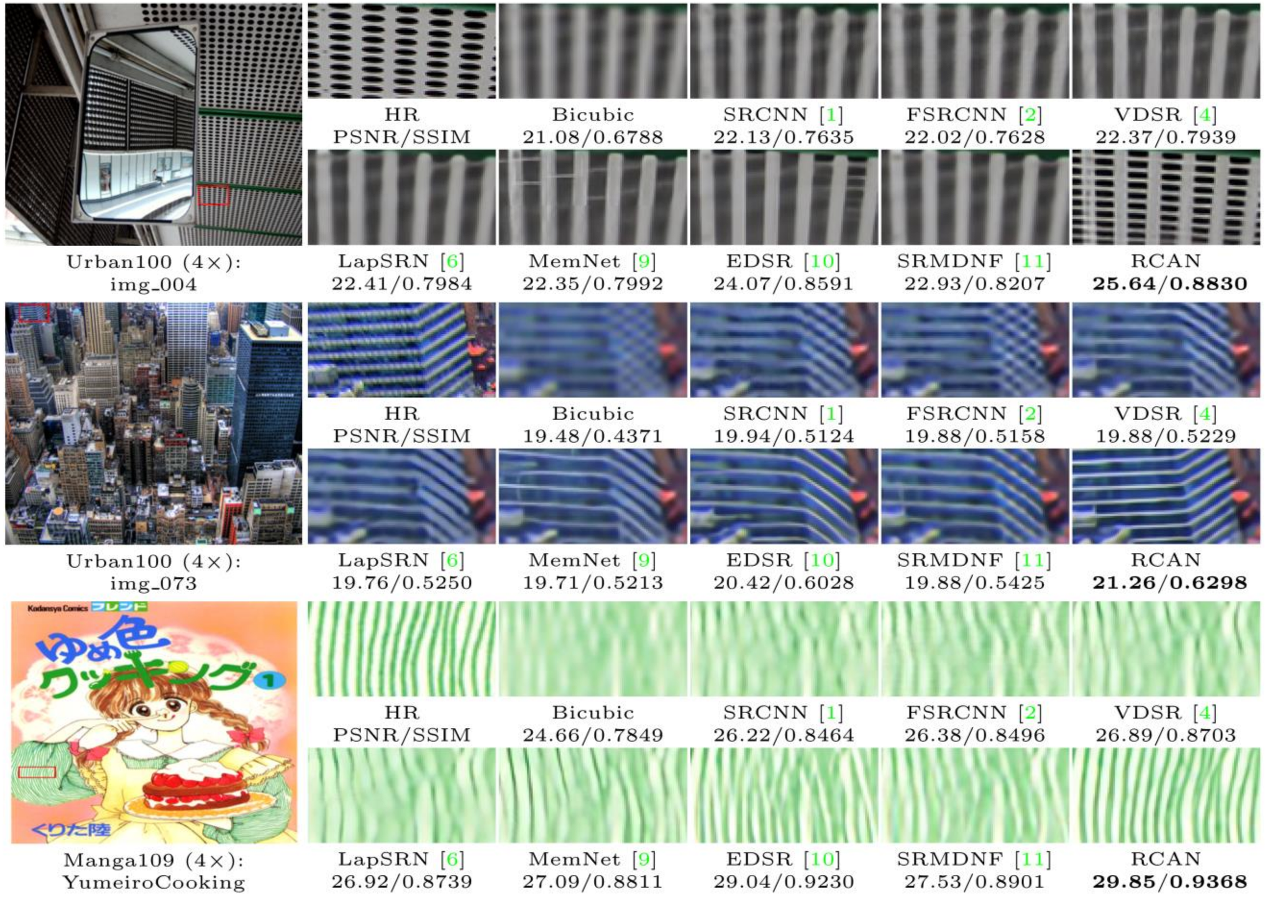


(a) Results on Set5 (4×) (b) Results on Set5 (8×)

Performance and number of parameters. Results are evaluated on Set5

Method	Scale	Set5 PSNR	Set5 SSIM	Set14 PSNR	Set14 SSIM	B100 PSNR	B100 SSIM	Urban100 PSNR	Urban100 SSIM	Manga109 PSNR	Manga109 SSIM
Bicubic	×2	33.66	0.9299	30.24	0.8688	29.56	0.8431	26.88	0.8403	30.80	0.9339
SRCNN [1]	×2	36.66	0.9542	32.45	0.9067	31.36	0.8879	29.50	0.8946	35.60	0.9663
FSRCNN [2]	×2	37.05	0.9560	32.66	0.9090	31.53	0.8920	29.88	0.9020	36.67	0.9710
VDSR [4]	×2	37.53	0.9590	33.05	0.9130	31.90	0.8960	30.77	0.9140	37.22	0.9750
LapSRN [6]	×2	37.52	0.9591	33.08	0.9130	31.08	0.8950	30.41	0.9101	37.27	0.9740
MemNet [9]	×2	37.78	0.9597	33.28	0.9142	32.08	0.8978	31.31	0.9195	37.72	0.9740
EDSR [10]	×2	38.11	0.9602	33.92	0.9195	32.32	0.9013	32.93	0.9351	39.10	0.9773
SRMDNF [11]	×2	37.79	0.9601	33.32	0.9159	32.05	0.8985	31.33	0.9204	38.07	0.9761
D-DBPN [16]	×2	38.09	0.9600	33.85	0.9190	32.27	0.9000	32.55	0.9324	38.89	0.9775
RDN [17]	×2	38.24	0.9614	34.01	0.9212	32.34	0.9017	32.89	0.9353	39.18	0.9780
RCAN (ours)	×2	38.27	0.9614	34.12	0.9216	32.41	0.9027	33.34	0.9384	39.44	0.9786
RCAN+ (ours)	×2	**38.33**	**0.9617**	**34.23**	**0.9225**	**32.46**	**0.9031**	**33.54**	**0.9399**	**39.61**	**0.9788**
Bicubic	×3	30.39	0.8682	27.55	0.7742	27.21	0.7385	24.46	0.7349	26.95	0.8556
SRCNN [1]	×3	32.75	0.9090	29.30	0.8215	28.41	0.7863	26.24	0.7989	30.48	0.9117
FSRCNN [2]	×3	33.18	0.9140	29.37	0.8240	28.53	0.7910	26.43	0.8080	31.10	0.9210
VDSR [4]	×3	33.67	0.9210	29.78	0.8320	28.83	0.7990	27.14	0.8290	32.01	0.9340
LapSRN [6]	×3	33.82	0.9227	29.87	0.8320	28.82	0.7980	27.07	0.8280	32.21	0.9350
MemNet [9]	×3	34.09	0.9248	30.00	0.8350	28.96	0.8001	27.56	0.8376	32.51	0.9369
EDSR [10]	×3	34.65	0.9280	30.52	0.8462	29.25	0.8093	28.80	0.8653	34.17	0.9476
SRMDNF [11]	×3	34.12	0.9254	30.04	0.8382	28.97	0.8025	27.57	0.8398	33.00	0.9403
RDN [17]	×3	34.71	0.9296	30.57	0.8468	29.26	0.8093	28.80	0.8653	34.13	0.9484
RCAN (ours)	×3	34.74	0.9299	30.65	0.8482	29.32	0.8111	29.09	0.8702	34.44	0.9499
RCAN+ (ours)	×3	**34.85**	**0.9305**	**30.76**	**0.8494**	**29.39**	**0.8122**	**29.31**	**0.8736**	**34.76**	**0.9513**
Bicubic	×4	28.42	0.8104	26.00	0.7027	25.96	0.6675	23.14	0.6577	24.89	0.7866
SRCNN [1]	×4	30.48	0.8628	27.50	0.7513	26.90	0.7101	24.52	0.7221	27.58	0.8555
FSRCNN [2]	×4	30.72	0.8660	27.61	0.7550	26.98	0.7150	24.62	0.7280	27.90	0.8610
VDSR [4]	×4	31.35	0.8830	28.02	0.7680	27.29	0.0726	25.18	0.7540	28.83	0.8870
LapSRN [6]	×4	31.54	0.8850	28.19	0.7720	27.32	0.7270	25.21	0.7560	29.09	0.8900
MemNet [9]	×4	31.74	0.8893	28.26	0.7723	27.40	0.7281	25.50	0.7630	29.42	0.8942
EDSR [10]	×4	32.46	0.8968	28.80	0.7876	27.71	0.7420	26.64	0.8033	31.02	0.9148
SRMDNF [11]	×4	31.96	0.8925	28.35	0.7787	27.49	0.7337	25.68	0.7731	30.09	0.9024
D-DBPN [16]	×4	32.47	0.8980	28.82	0.7860	27.72	0.7400	26.38	0.7946	30.91	0.9137
RDN [17]	×4	32.47	0.8990	28.81	0.7871	27.72	0.7419	26.61	0.8028	31.00	0.9151
RCAN (ours)	×4	32.63	0.9002	28.87	0.7889	27.77	0.7436	26.82	0.8087	31.22	0.9173
RCAN+ (ours)	×4	**32.73**	**0.9013**	**28.98**	**0.7910**	**27.85**	**0.7455**	**27.10**	**0.8142**	**31.65**	**0.9208**
Bicubic	×8	24.40	0.6580	23.10	0.5660	23.67	0.5480	20.74	0.5160	21.47	0.6500
SRCNN [1]	×8	25.33	0.6900	23.76	0.5910	24.13	0.5660	21.29	0.5440	22.46	0.6950
FSRCNN [2]	×8	20.13	0.5520	19.75	0.4820	24.21	0.5680	21.32	0.5380	22.39	0.6730
SCN [3]	×8	25.59	0.7071	24.02	0.6028	24.30	0.5698	21.52	0.5571	22.68	0.6963
VDSR [4]	×8	25.93	0.7240	24.26	0.6140	24.49	0.5830	21.70	0.5710	23.16	0.7250
LapSRN [6]	×8	26.15	0.7380	24.35	0.6200	24.54	0.5860	21.81	0.5810	23.39	0.7350
MemNet [9]	×8	26.16	0.7414	24.38	0.6199	24.58	0.5842	21.89	0.5825	23.56	0.7387
MSLapSRN [7]	×8	26.34	0.7558	24.57	0.6273	24.65	0.5895	22.06	0.5963	23.90	0.7564
EDSR [10]	×8	26.96	0.7762	24.91	0.6420	24.81	0.5985	22.51	0.6221	24.69	0.7841
D-DBPN [16]	×8	27.21	0.7840	25.13	0.6480	24.88	0.6010	22.73	0.6312	25.14	0.7987
RCAN (ours)	×8	27.31	0.7878	25.23	0.6511	24.98	0.6058	23.00	0.6452	25.24	0.8029
RCAN+ (ours)	×8	**27.47**	**0.7913**	**25.40**	**0.6553**	**25.05**	**0.6077**	**23.22**	**0.6524**	**25.58**	**0.8092**

Experiment results



Conclusion

- Main contribution:
  - Residual in residual (RIR) structure Using long and short skip connection
    → very deep trainable networks, learn more effective information
  - Channel attention mechanism → improves the representational ability

→highly accurate image SR