RCAN

Image Super-Resolution Using Very Deep Residual Channel Attention Networks

Dajin Han

Index

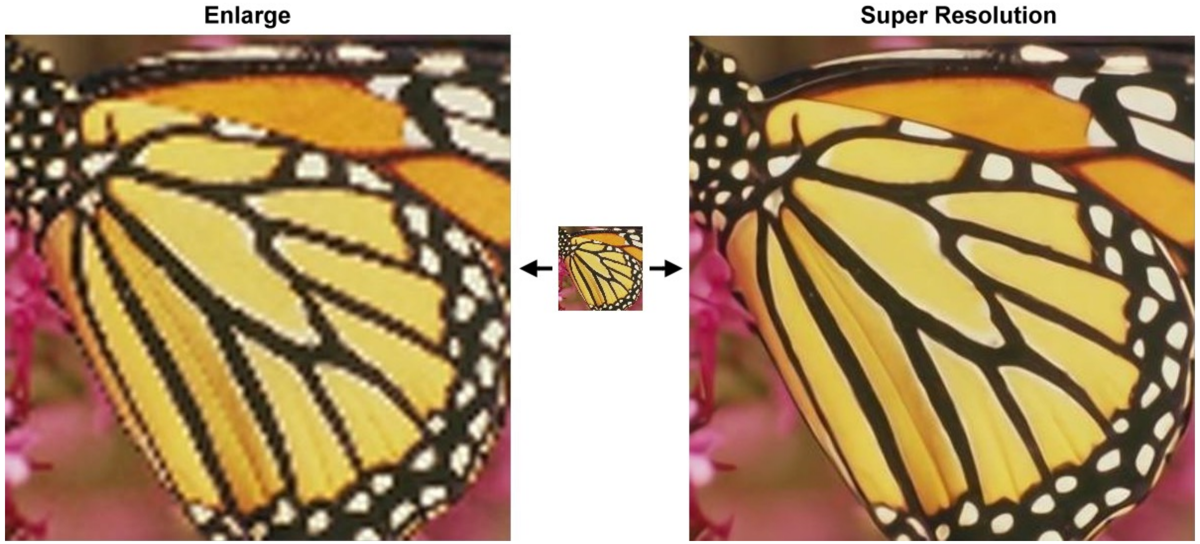
- Prior studies
- Architecture
  - Residual in Residual
  - Channel Attention
  - Residual Channel Attention Blocks
- Experiments
- QA

Prior studies

Concepts, SRCNN, VDSR, SRGAN, EDSR, MDSR

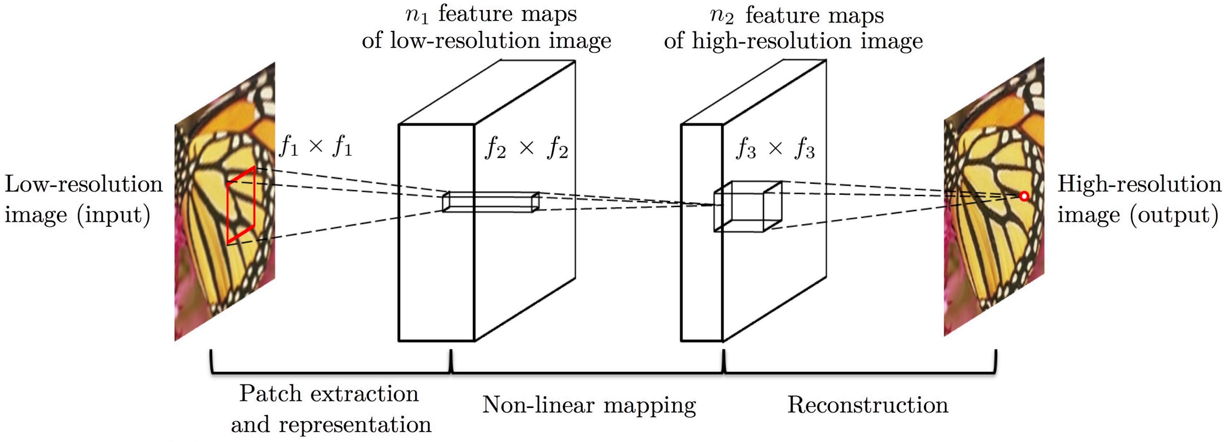
Concept of Super Resolution

- LR to HR
- Methodology
  - Local prior
    - interpolation
    - cnn-based
  - Non-local prior
    - patch
    - cnn-based



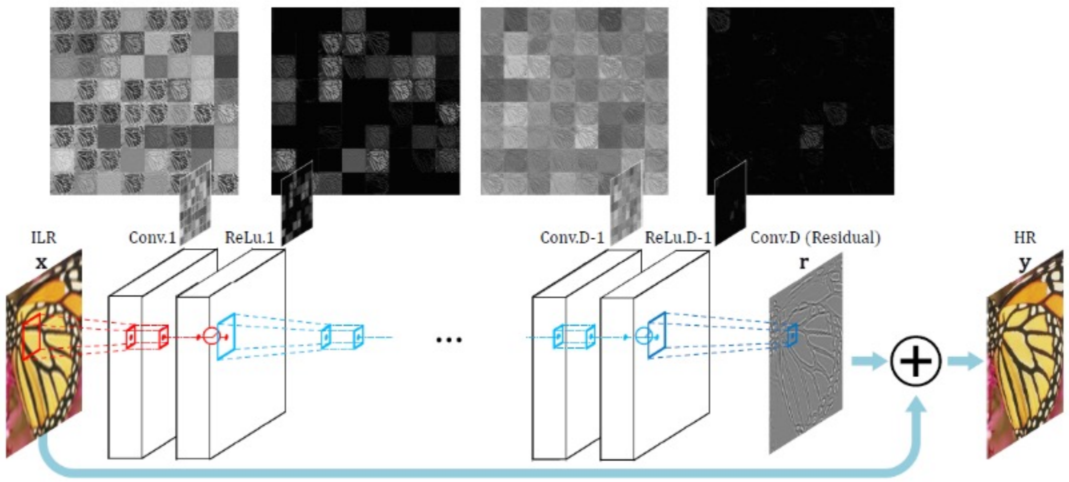
SRCNN

1. Upscale
2. LR feature map
3. Non-linear mapping
4. HR feature map
5. Reconstruction, MSE



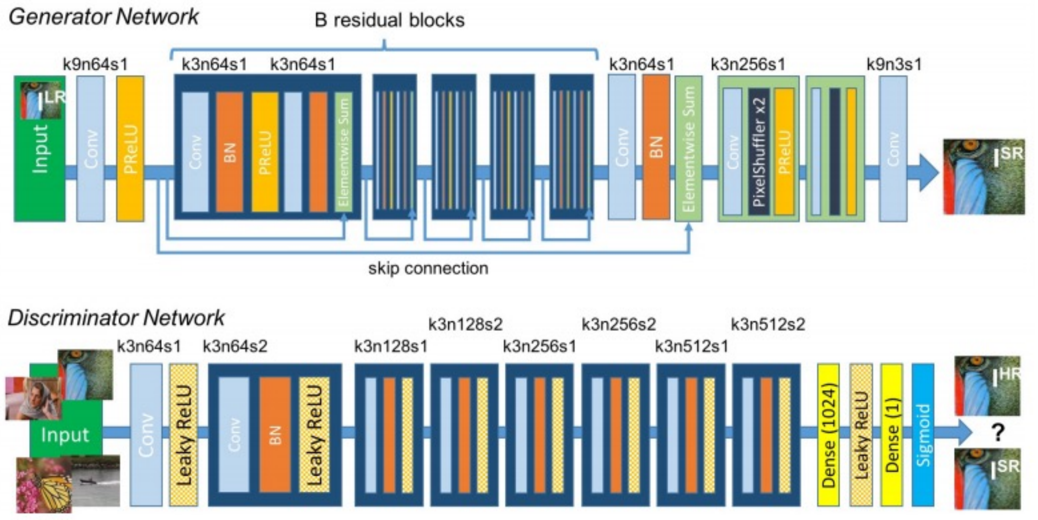
VDSR

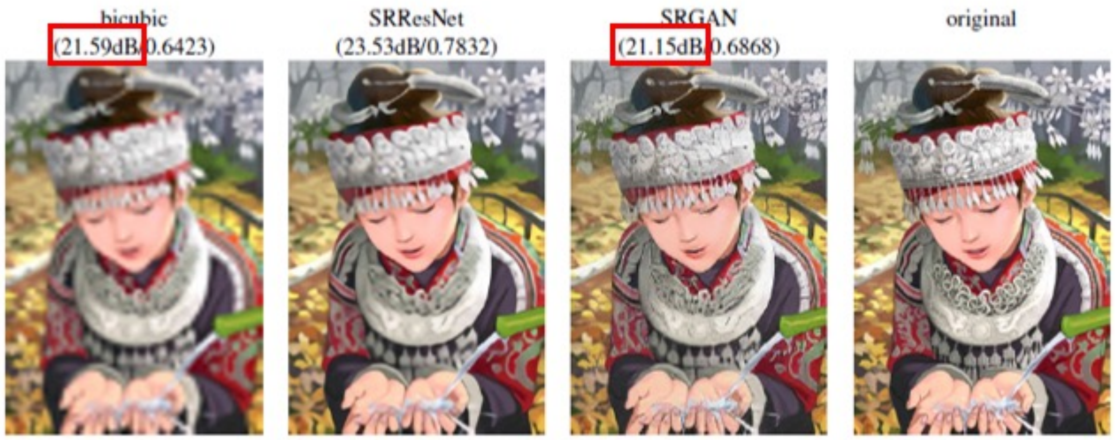
- Deep. Inspired by VGG-net
- More contextual info
- “Residual” connection
- MSE



SRGAN

- GAN methodology
- Perceptual loss
- Less blur





EDSR, MDSR

- Advanced residual block
  - Remove BN layer
  - Less normalize, more detail
- L1 loss

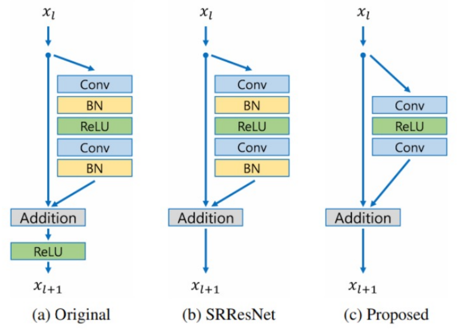


Fig4. comparison of residual blocks

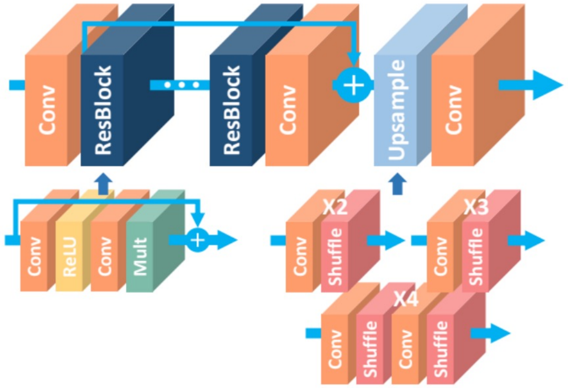


Fig5. EDSR(single-scale)architecture

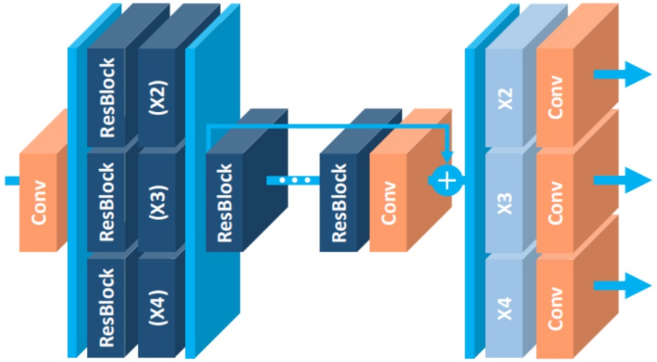


Fig6. MDSR(multi-scale) architecture

RCAN

- SR model
- **Residual in Residual**
- **Channel Attention**
- Rescale feature via intra-channel information
- L1 loss

Architecture

RIR, CA, RCAB

Architecture

Yulun Zhang *et al.*

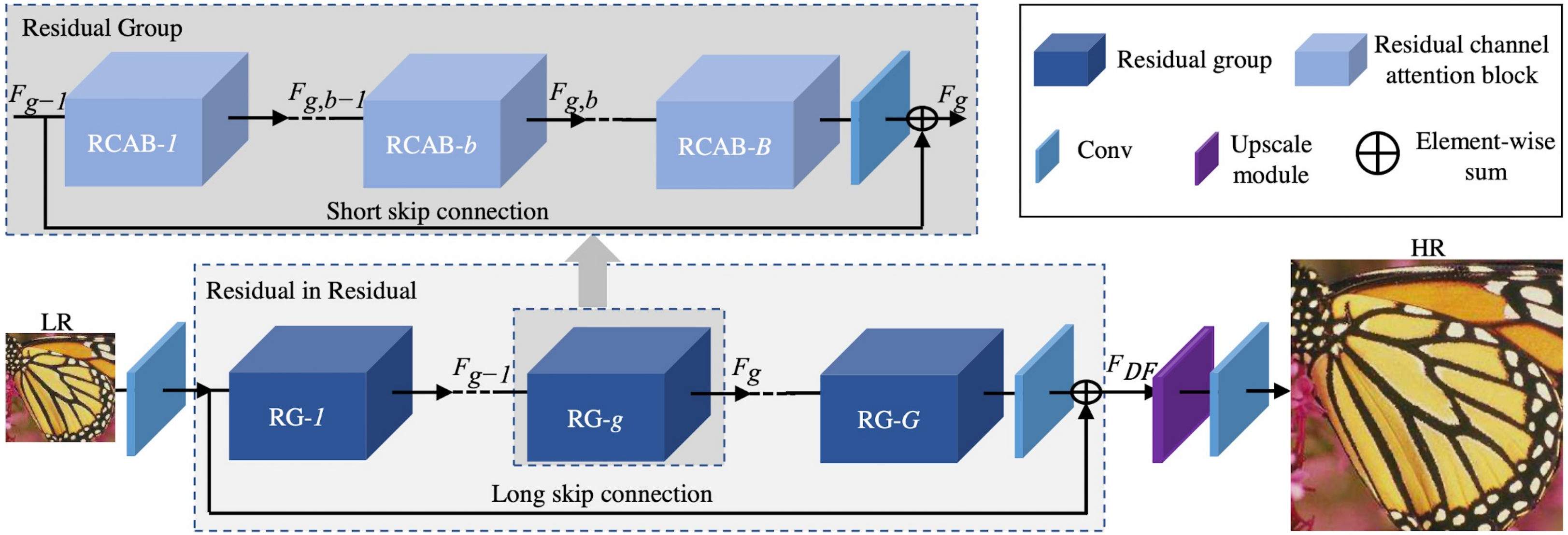


Fig. 2. Network architecture of our residual channel attention network (RCAN)

Architecture

1. **Shallow feature extraction**
2. RIR deep feature extraction
3. Upscale module
4. Reconstruction

Yulun Zhang *et al.*

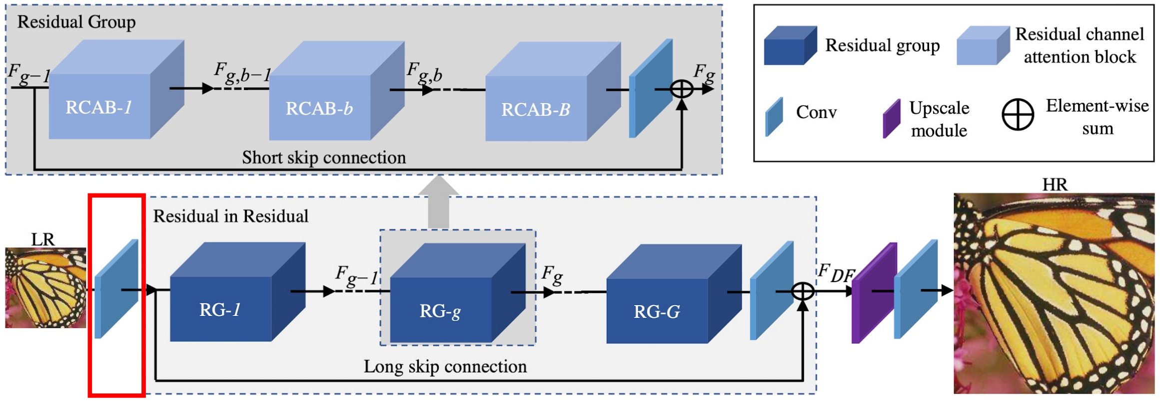


Fig. 2. Network architecture of our residual channel attention network (RCAN)

$$F_0 = H_{SF}(I_{LR}),$$

I_{LR} : Input LR Image
H_{SF} : Shallow Feature Extraction Layer
F_0 : Shallow Feature

Architecture

1. Shallow feature extraction
2. **RIR deep feature extraction**
3. Upscale module
4. Reconstruction

Yulun Zhang *et al.*

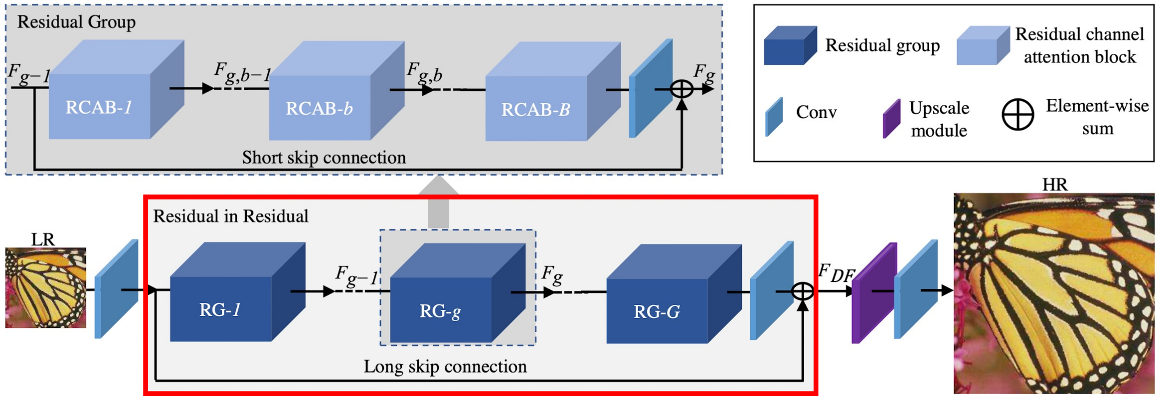


Fig. 2. Network architecture of our residual channel attention network (RCAN)

$$F_{DF} = H_{RIR}(F_0),$$

F_0 : Input shallow feature
H_{RIR} : RIR Layer
F_{DF} : Deep Feature

Architecture

1. Shallow feature extraction
2. RIR deep feature extraction
3. **Upscale module**
4. Reconstruction

Yulun Zhang *et al.*

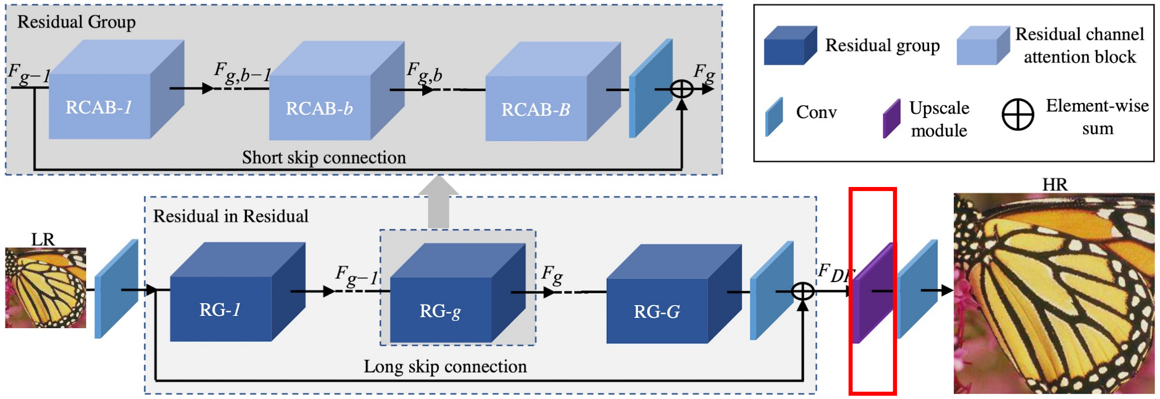


Fig. 2. Network architecture of our residual channel attention network (RCAN)

$$F_{UP} = H_{UP}\left(F_{DF}\right),$$

F_{DF} : Deep Feature
H_{UP} : Upsacle Layer
F_{UP} : Upscaled Feature

Architecture

1. Shallow feature extraction
2. RIR deep feature extraction
3. Upscale module
4. **Reconstruction**

Yulun Zhang *et al.*

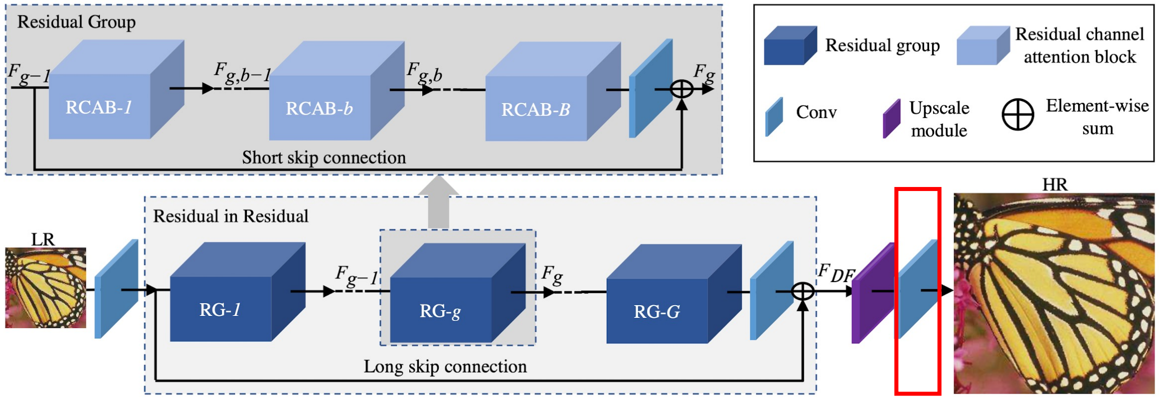


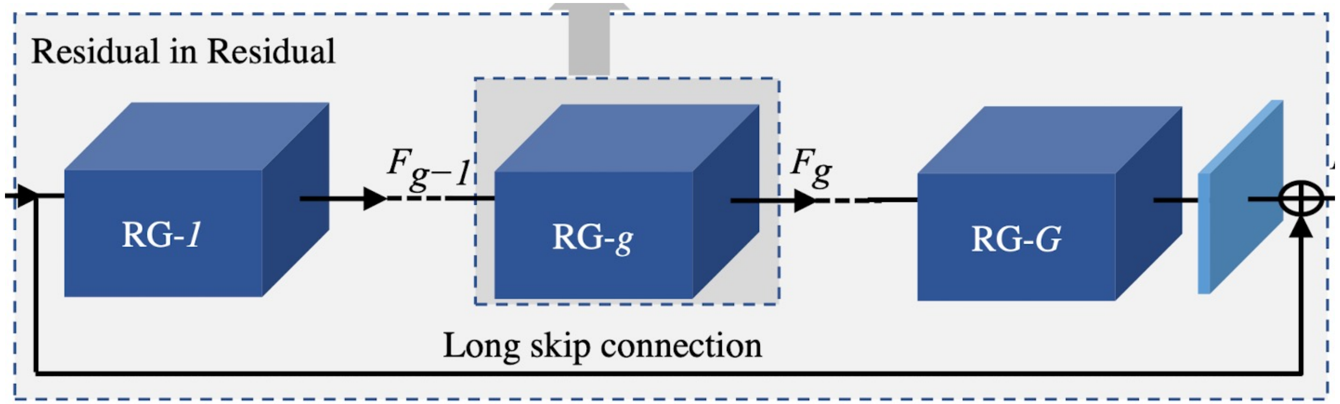
Fig. 2. Network architecture of our residual channel attention network (RCAN)

$$I_{SR} = H_{REC}(F_{UP}) = H_{RCAN}(I_{LR})$$

F_{UP} : Upscaled Feature
H_{REC} : Reconstruction Layer
I_{SR} : SR Image
H_{RCAN} : RCAN model

Residual in Residual

- G Residual Groups (RG)
- Long skip connection (LSC)
- Convolution at tail of RIR



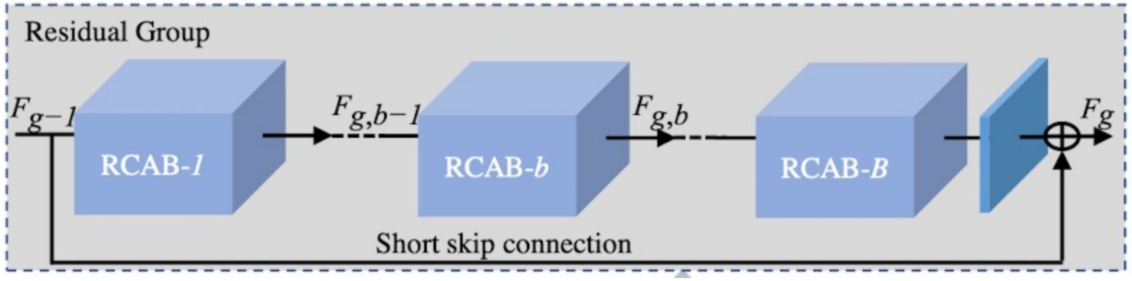
$$F_g = H_g(F_{g-1}) \\ = H_g(H_{g-1}(\cdots H_1(F_0)\cdots))$$

F_g : feature, output of g_th Residual Group
H_g : g_th Residual Group
F_{DF} : Deep Feature
W_{LSC} : Weight of Convolution Layer

$$F_{DF} = F_0 + W_{LSC}F_G \\ = F_0 + W_{LSC}H_g(H_{g-1}(\cdots H_1(F_0)\cdots))$$

Residual Group

- B Residual Channel Attention Blocks (RCAB)
- Convolution at tail of RG
- Short skip connection (SSC)



$$
\begin{aligned}
F_{g,b} &= H_{g,b}\left(F_{g,b-1}\right) \\
&= H_{g,b}\left(H_{g,b-1}\left(\cdots H_{g,1}\left(F_{g-1}\right)\cdots\right)\right)
\end{aligned}
$$

$$
\begin{aligned}
F_g &= F_{g-1} + W_g F_{g,B} \\
&= F_{g-1} + W_g H_{g,B}\left(H_{g,B-1}\left(\cdots H_{g,1}\left(F_{g-1}\right)\cdots\right)\right)
\end{aligned}
$$

$F_{g,b}$: feature, output of g_th RG, b_th RCAB
$H_{g,b}$: b_th RCAB of g_th RG
W_g : Weight of Convolution Layer of g_th RG

Channel Attention

- High frequency : texture, edge
- Low frequency : complanate
- Conv layer
  - Unable to exploit contextual information outside of the local region
  - Unable to get region information of high frequency

Channel Attention

- Global pooling
  - Contextual Information outside of the local region

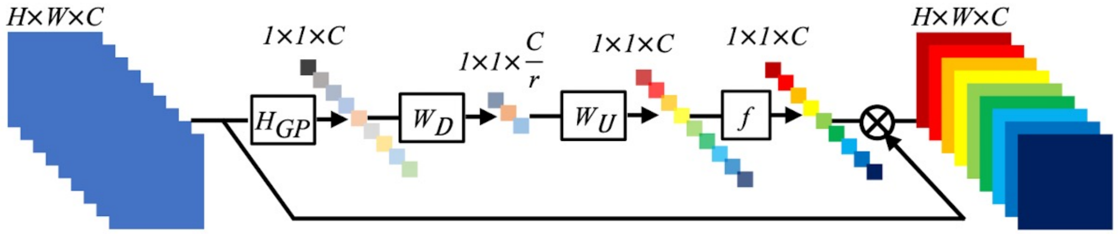


Fig. 3. Channel attention (CA). ⊗ denotes element-wise product

$$X = [x_1, \cdots, x_c, \cdots, x_C]$$

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

$$s = f(W_U \delta(W_D z))$$

$$\widehat{x}_c = s_c \cdot x_c,$$

X : Input Feature
$x_c(i,j)$: c_th channel, i_th row, j_th column value of X
H_{GP} : Global Pooling layer
W_D , W_U : Weights of convolution layers
δ : ReLU
f : Sigmoid

Residual Channel Attention Block

$$F_{g,b} = F_{g,b-1} + R_{g,b}\left(X_{g,b}\right) \cdot X_{g,b},$$

$$X_{g,b} = W_{g,b}^{2} \delta\left(W_{g,b}^{1} F_{g,b-1}\right),$$

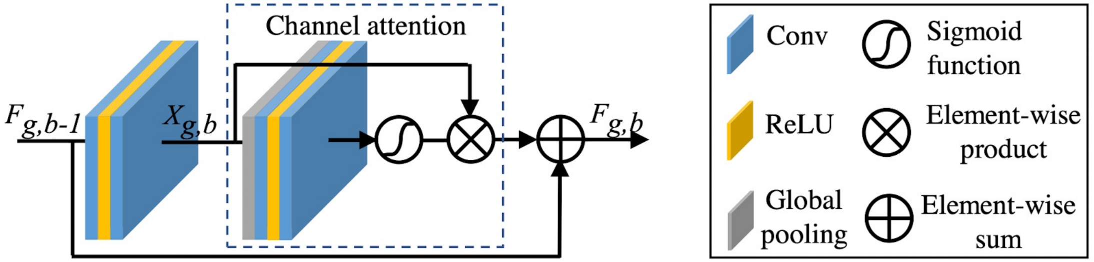


Fig. 4. Residual channel attention block (RCAB)

Table 1. Investigations of RIR (including LSC and SSC) and CA. We observe the best PSNR (dB) values on Set5 ($2\times$) in 5×10^4 iterations

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

X : Input Feature
$F_{g\,b}$: feature, output of g_th RG, b_th RCAB
$R_{g,b}$: Channel attention of g_th RG, b_th RCAB
$W_{g,b}^{n}$: Weights of convolution layer in g_th RG, b_th RCAB
δ : ReLU

Loss

1. Shallow feature extraction
2. RIR deep feature extraction
3. Upscale module
4. **Reconstruction**

Yulun Zhang *et al.*

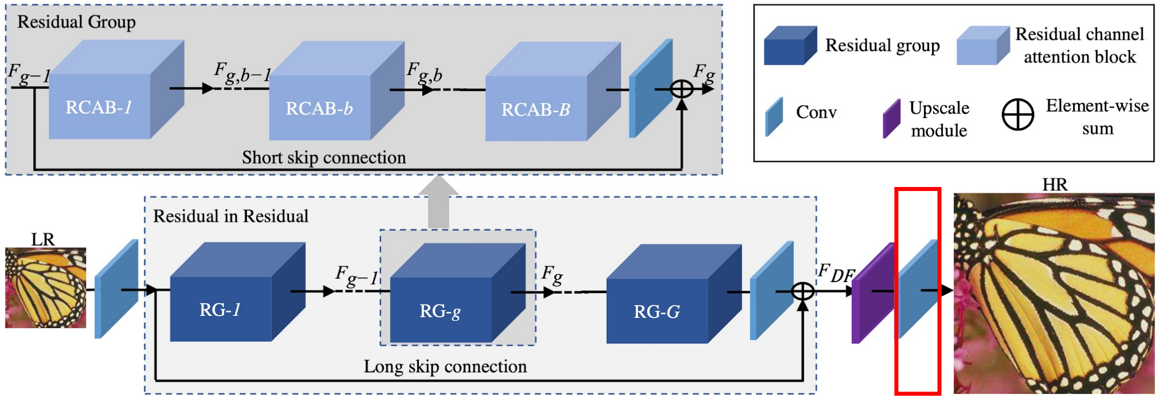


Fig. 2. Network architecture of our residual channel attention network (RCAN)

$$L(\Theta) = \frac{1}{N}\sum_{i=1}^{N}\left\|H_{RCAN}\left(I_{LR}^{i}\right) - I_{HR}^{i}\right\|_1 ,$$

Experiments

Experiments

Table 2. Quantitative results with BI degradation model. Best and second best results are **highlighted** and underlined

Method	Scale	Set5		Set14		B100		Urban100		Manga109	
		PSNR	SSIM	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	30.80	0.9339
SRCNN [5]	×2	36.66	0.9542	32.45	0.9067	31.36	0.8879	29.50	0.8946	35.60	0.9663
FSRCNN [6]	×2	37.05	0.9560	32.66	0.9090	31.53	0.8920	29.88	0.9020	36.67	0.9710
VDSR [16]	×2	37.53	0.9590	33.05	0.9130	31.90	0.8960	30.77	0.9140	37.22	0.9750
LapSRN [19]	×2	37.52	0.9591	33.08	0.9130	31.08	0.8950	30.41	0.9101	37.27	0.9740
MemNet [35]	×2	37.78	0.9597	33.28	0.9142	32.08	0.8978	31.31	0.9195	37.72	0.9740
EDSR [23]	×2	38.11	0.9602	33.92	0.9195	32.32	0.9013	32.93	0.9351	39.10	0.9773
SRMDNF [43]	×2	37.79	0.9601	33.32	0.9159	32.05	0.8985	31.33	0.9204	38.07	0.9761
D-DBPN [10]	×2	38.09	0.9600	33.85	0.9190	32.27	0.9000	32.55	0.9324	38.89	0.9775
RDN [44]	×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 [5]	×3	32.75	0.9090	29.30	0.8215	28.41	0.7863	26.24	0.7989	30.48	0.9117
FSRCNN [6]	×3	33.18	0.9140	29.37	0.8240	28.53	0.7910	26.43	0.8080	31.10	0.9210
VDSR [16]	×3	33.67	0.9210	29.78	0.8320	28.83	0.7990	27.14	0.8290	32.01	0.9340
LapSRN [19]	×3	33.82	0.9227	29.87	0.8320	28.82	0.7980	27.07	0.8280	32.21	0.9350
MemNet [35]	×3	34.09	0.9248	30.00	0.8350	28.96	0.8001	27.56	0.8376	32.51	0.9369
EDSR [23]	×3	34.65	0.9280	30.52	0.8462	29.25	0.8093	28.80	0.8653	34.17	0.9476
SRMDNF [43]	×3	34.12	0.9254	30.04	0.8382	28.97	0.8025	27.57	0.8398	33.00	0.9403
RDN [44]	×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 [5]	×4	30.48	0.8628	27.50	0.7513	26.90	0.7101	24.52	0.7221	27.58	0.8555
FSRCNN [6]	×4	30.72	0.8660	27.61	0.7550	26.98	0.7150	24.62	0.7280	27.90	0.8610
VDSR [16]	×4	31.35	0.8830	28.02	0.7680	27.29	0.0726	25.18	0.7540	28.83	0.8870
LapSRN [19]	×4	31.54	0.8850	28.19	0.7720	27.32	0.7270	25.21	0.7560	29.09	0.8900
MemNet [35]	×4	31.74	0.8893	28.26	0.7723	27.40	0.7281	25.50	0.7630	29.42	0.8942
EDSR [23]	×4	32.46	0.8968	28.80	0.7876	27.71	0.7420	26.64	0.8033	31.02	0.9148
SRMDNF [43]	×4	31.96	0.8925	28.35	0.7787	27.49	0.7337	25.68	0.7731	30.09	0.9024
D-DBPN [10]	×4	32.47	0.8980	28.82	0.7860	27.72	0.7400	26.38	0.7946	30.91	0.9137
RDN [44]	×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 [5]	×8	25.33	0.6900	23.76	0.5910	24.13	0.5660	21.29	0.5440	22.46	0.6950
FSRCNN [6]	×8	20.13	0.5520	19.75	0.4820	24.21	0.5680	21.32	0.5380	22.39	0.6730
SCN [39]	×8	25.59	0.7071	24.02	0.6028	24.30	0.5698	21.52	0.5571	22.68	0.6963
VDSR [16]	×8	25.93	0.7240	24.26	0.6140	24.49	0.5830	21.70	0.5710	23.16	0.7250
LapSRN [19]	×8	26.15	0.7380	24.35	0.6200	24.54	0.5860	21.81	0.5810	23.39	0.7350
MemNet [35]	×8	26.16	0.7414	24.38	0.6199	24.58	0.5842	21.89	0.5825	23.56	0.7387
MSLapSRN [20]	×8	26.34	0.7558	24.57	0.6273	24.65	0.5895	22.06	0.5963	23.90	0.7564
EDSR [23]	×8	26.96	0.7762	24.91	0.6420	24.81	0.5985	22.51	0.6221	24.69	0.7841
D-DBPN [10]	×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**

Experiments

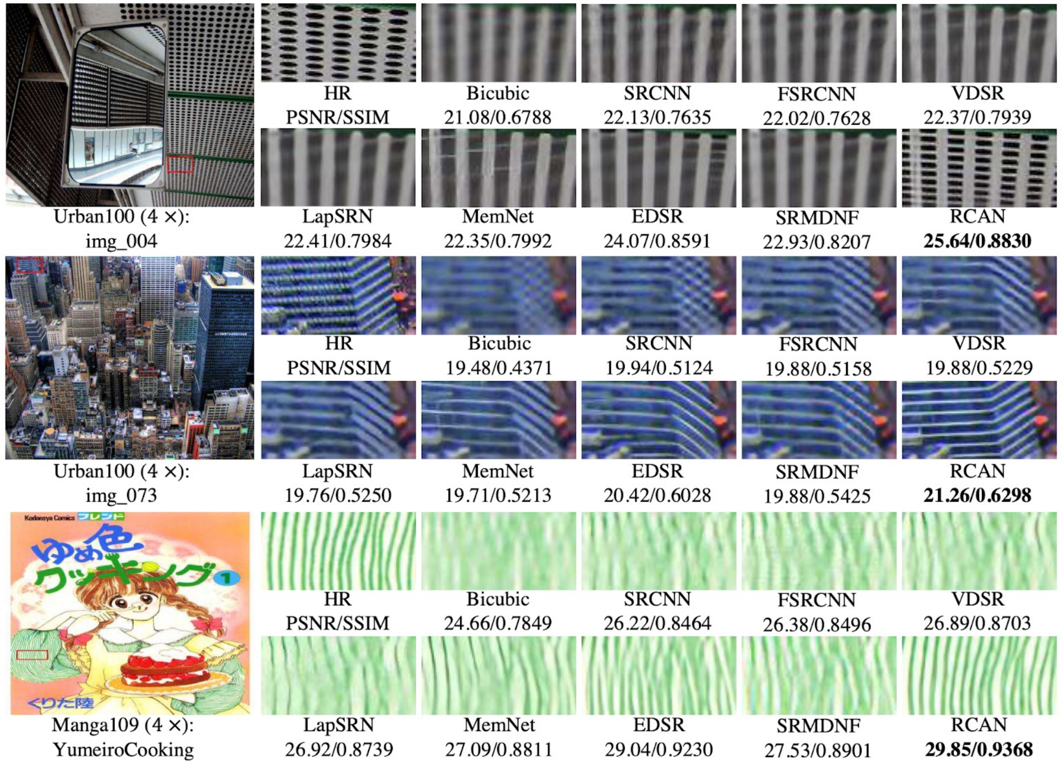


Fig. 5. Visual comparison for 4× SR with BI model on Urban100 and Manga109 datasets. The best results are **highlighted**

Experiments

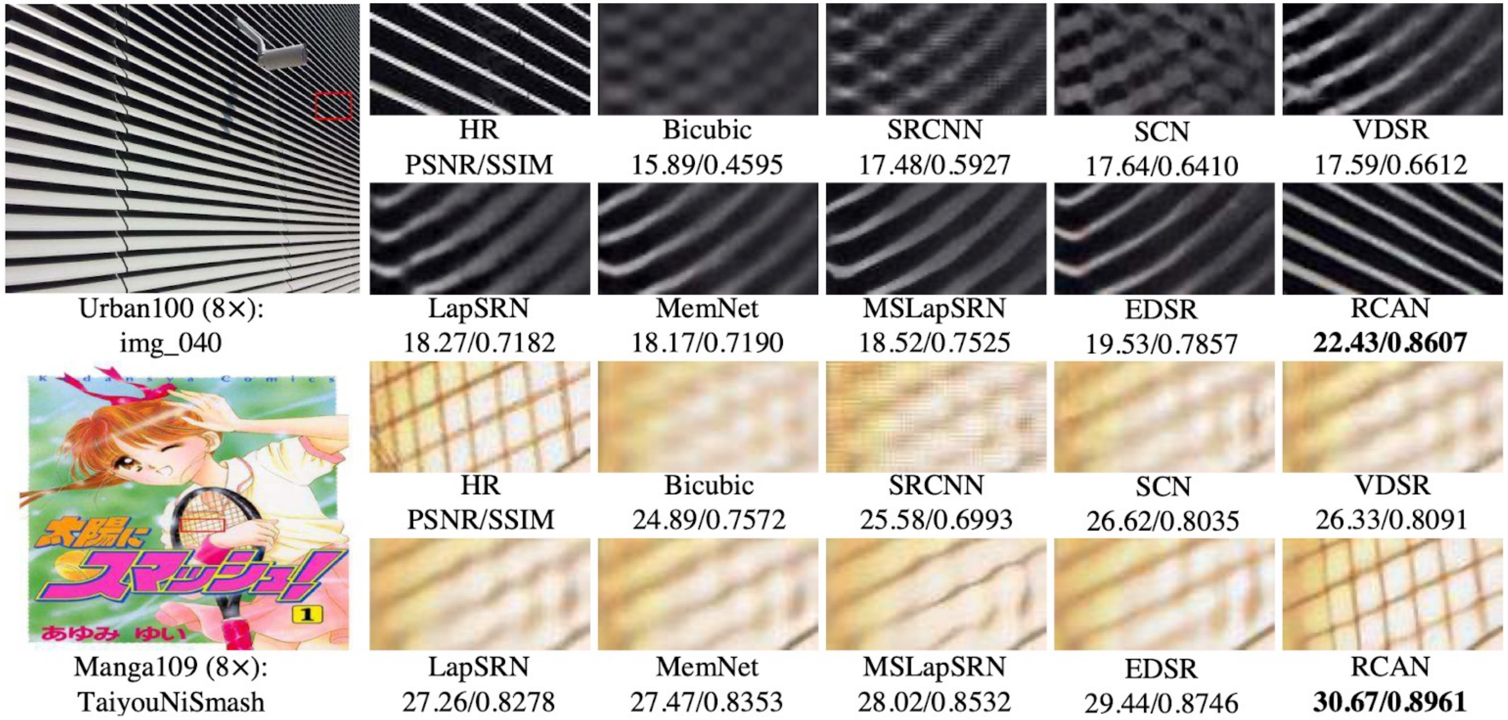


Fig. 6. Visual comparison for 8× SR with BI model on Urban100 and Manga109 datasets. The best results are **highlighted**

Experiments

Table 3. Quantitative results with BD degradation model. Best and second best results are **highlighted** and underlined

Method	Scale	Set5		Set14		B100		Urban100		Manga109	
		PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	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 [29]	×3	32.21	0.9001	28.89	0.8105	28.13	0.7740	25.84	0.7856	29.64	0.9003
SRCNN [5]	×3	32.05	0.8944	28.80	0.8074	28.13	0.7736	25.70	0.7770	29.47	0.8924
FSRCNN [6]	×3	26.23	0.8124	24.44	0.7106	24.86	0.6832	22.04	0.6745	23.04	0.7927
VDSR [16]	×3	33.25	0.9150	29.46	0.8244	28.57	0.7893	26.61	0.8136	31.06	0.9234
IRCNN [42]	×3	33.38	0.9182	29.63	0.8281	28.65	0.7922	26.77	0.8154	31.15	0.9245
SRMDNF [43]	×3	34.01	0.9242	30.11	0.8364	28.98	0.8009	27.50	0.8370	32.97	0.9391
RDN [44]	×3	34.58	0.9280	30.53	0.8447	29.23	0.8079	28.46	0.8582	33.97	0.9465
RCAN (ours)	×3	34.70	0.9288	30.63	0.8462	29.32	0.8093	28.81	0.8647	34.38	0.9483
RCAN+ (ours)	×3	**34.83**	**0.9296**	**30.76**	**0.8479**	**29.39**	**0.8106**	**29.04**	**0.8682**	**34.76**	**0.9502**

Experiments

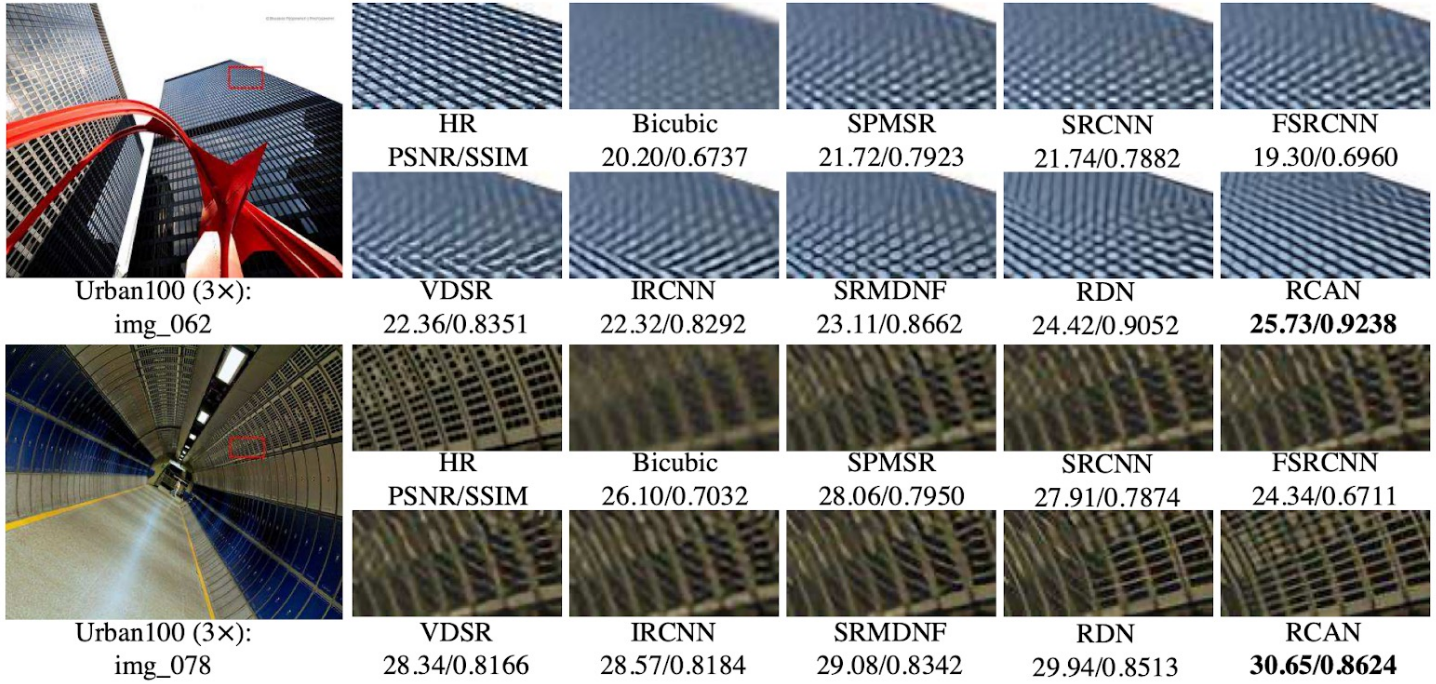


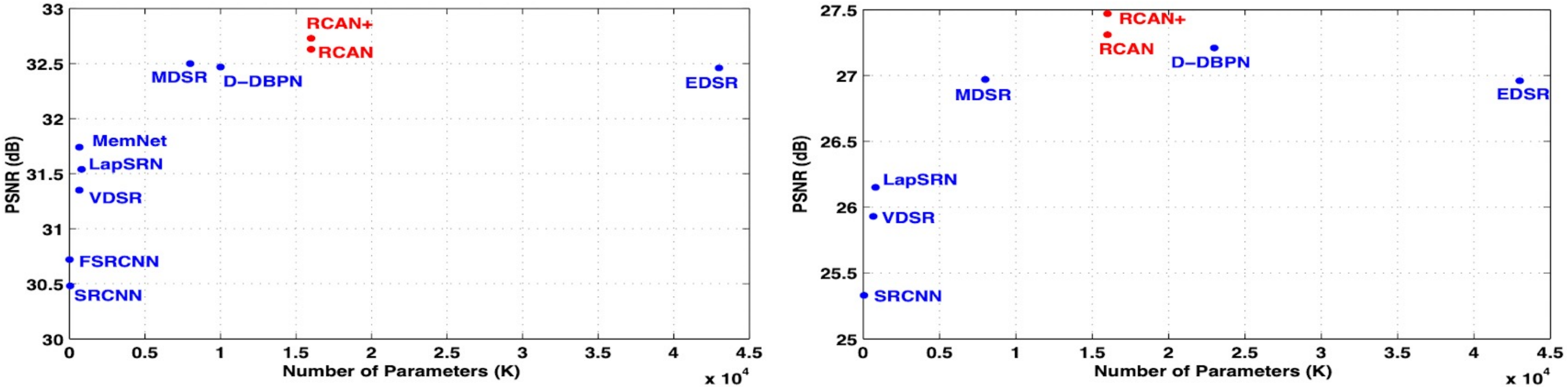
Fig. 7. Visual comparison for 3× SR with BD model on Urban100 dataset. The best results are **highlighted**

Experiments

Table 4. ResNet object recognition performance. The best results are **highlighted**

Evaluation	Bicubic	DRCN [17]	FSRCNN [6]	PSyCo [30]	ENet-E [31]	RCAN	Baseline
Top-1 error	0.506	0.477	0.437	0.454	0.449	**0.393**	0.260
Top-5 error	0.266	0.242	0.196	0.224	0.214	**0.167**	0.072

Experiments



(a) Results on Set5 (4×)

(b) Results on Set5 (8×)

Fig. 8. Performance and number of parameters. Results are evaluated on Set5

QA