Efficient Single Image Super-Resolution via Hybrid Residual Feature Learning with Compact Back-Projection Network

Feiyang Zhu, Qijun Zhao
College of Computer Science, Sichuan University, Chengdu, Sichuan, P. R. China

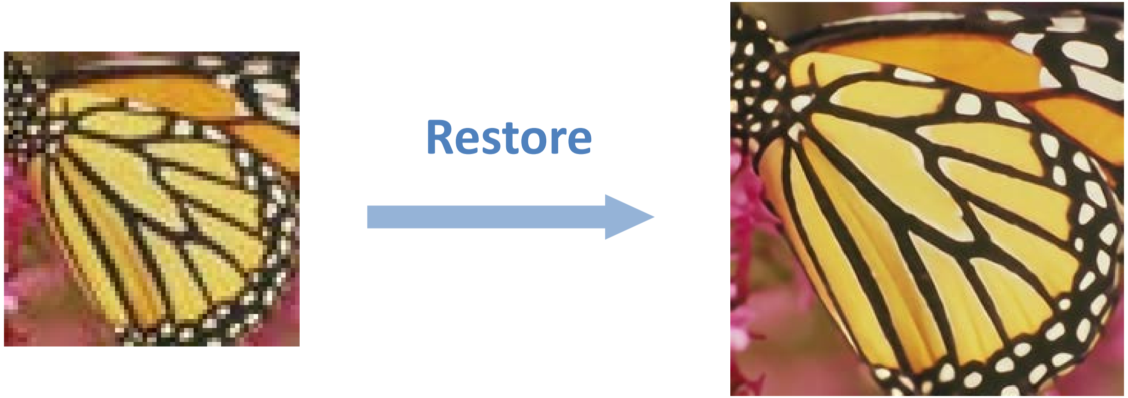
ICCV 2019, Seoul, Republic of Korea

Outline

- Background
- Related Work
- Method
- Experiment
- Conclusion

Background

- Single image super-resolution (SISR)

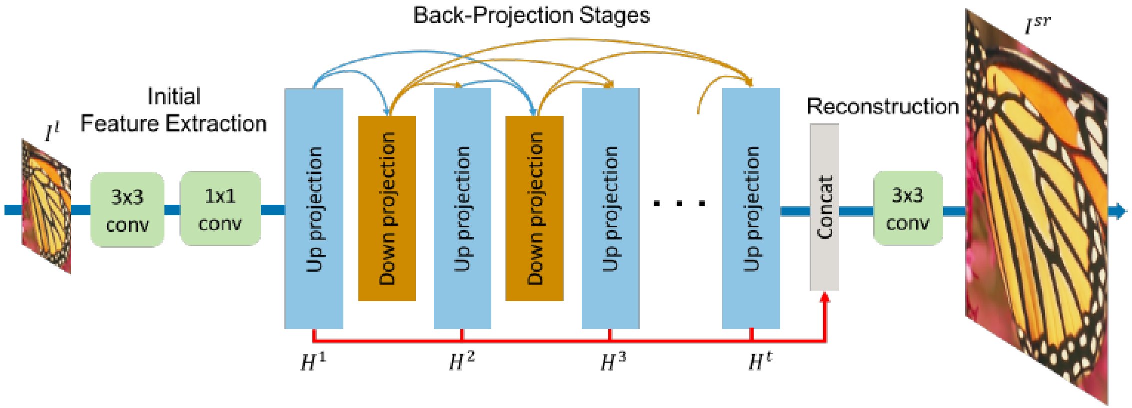


Low-resolution(LR) and lack of details   High-resolution(HR) and clear

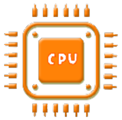
Related Work

- Deep Back-Projection Networks (DBPN[1])

Not practical



Parameters: 10,426K
Model size: 39.8MB



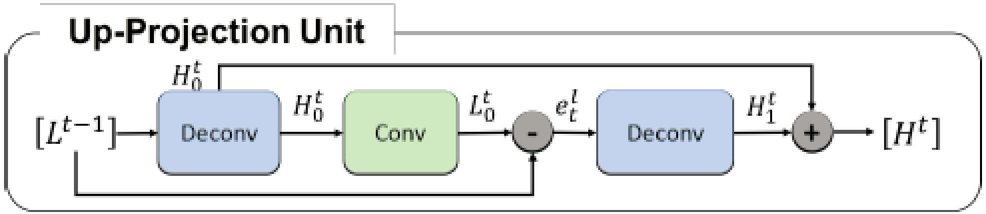
Mult-Adds: 5,112G
Time: 6.25s

- Iteration of up-projection unit and down-projection unit.
- Concat the output of the up-projection units to reconstruct the HR image.

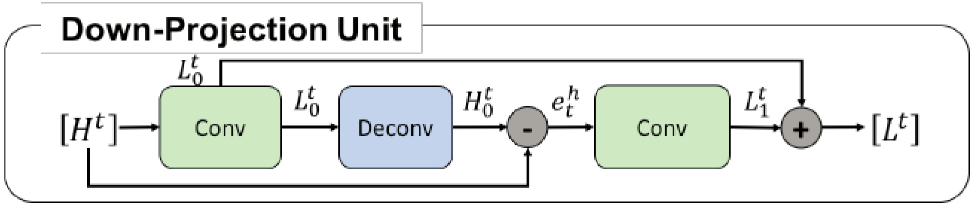
[1] M. Haris, G. Shakhnarovich, and N. Ukita. Deep back-projection networks for super-resolution. In CVPR, pages 1664–1673, 2018.

Related Work

- Projection unit of DBPN



Iteration of LR to HR features



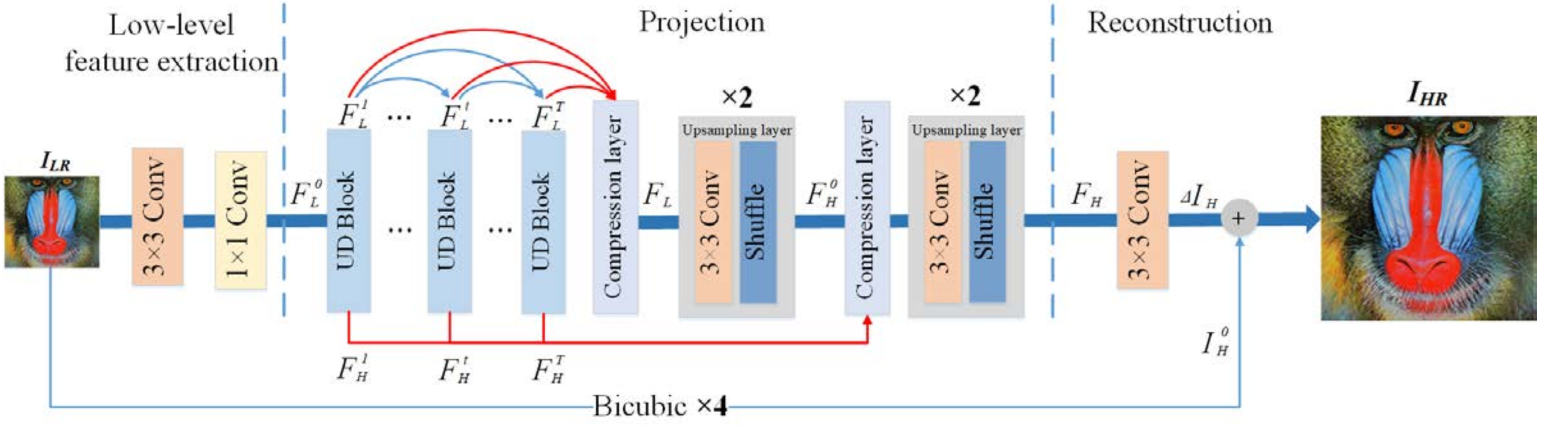
Iteration of HR to LR features

Method

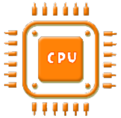
- Motivation for improving DBPN
  - The feature information in DBPN network is not fully utilized.
  - The large number of parameters and operations of DBPN.
  - The learning pressure of the network is too great.

Method

- Network structure on ×4 scale



Three parts： 1) low-level feature extraction
2) projection
3) reconstruction

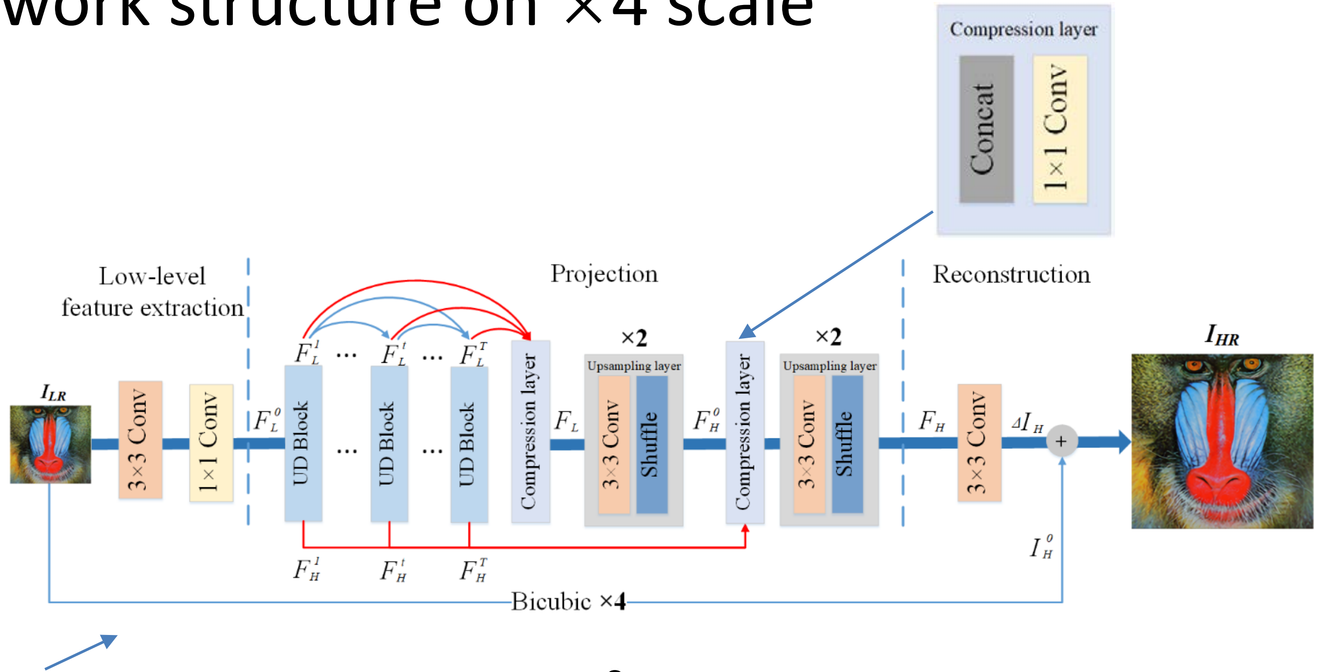


Parameters: 1,197K
Model size: 4.8MB

Mult-Adds: 97.9G
Time: 0.40s

Method

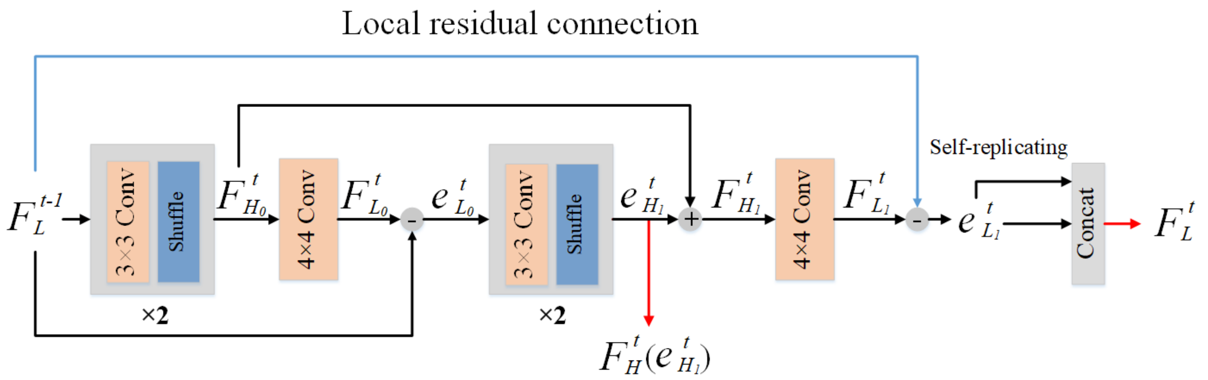
- Network structure on ×4 scale



Global residual connection $I_{HR} = I_H^0 + \Delta I_H$

Method

- Network structure: UD Block



- Deconvolution layer
- 64 filters
- Use HR features

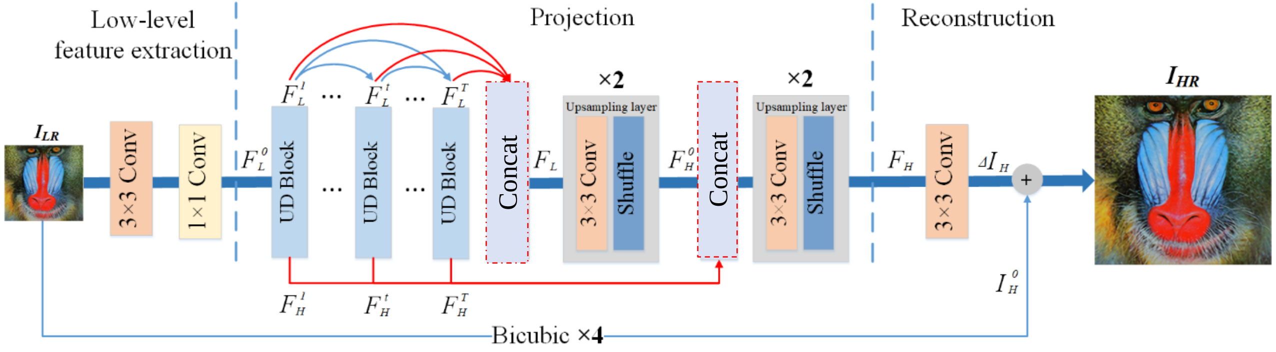
DBPN

- Sub-pixel convolution layer
- 32 filters
- Use hybrid residual features

Ours

Method

- More lightweight network



	Parameters	Mult-Adds
No compression layer Add compression layers	1,664K 1,197K	122.0G 97.9G
64 filters 32 filters	4,746K 1,197K	388.0G 97.9G

Experiments

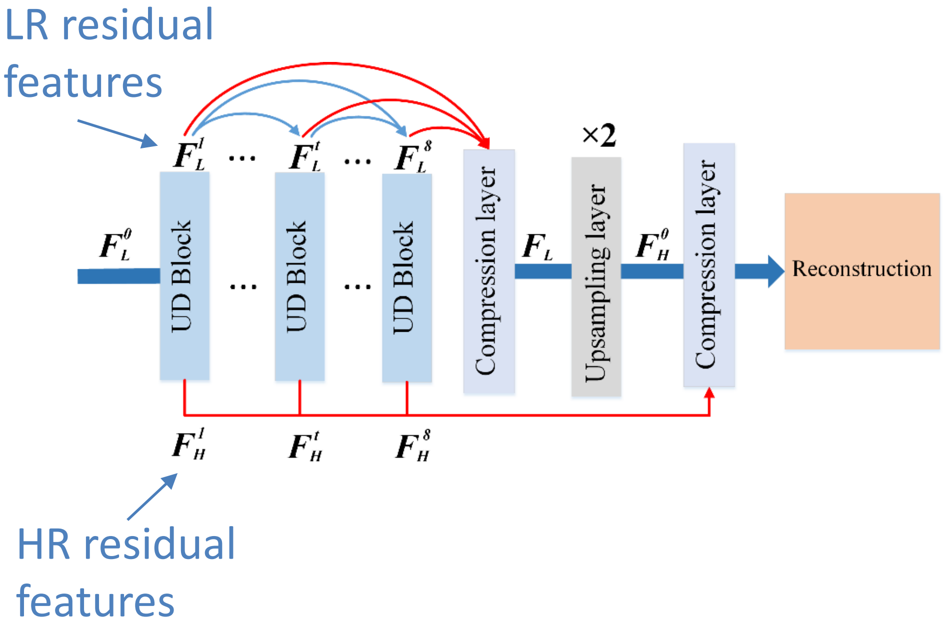
- Dataset
  - DIV2K(800 images for training, 100 images for validating)
  - Testing dataset: Set5/Set14/BSDS100/Urban100

- Data expansion
  - Randomly flipping LR images horizontally or vertically
  - Randomly rotating LR images by 90°

Experiments

What features are better in the reconstruction?

LR or HR residual features ? or hybrid residual features?

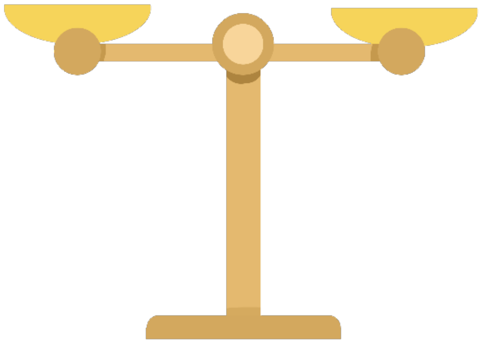


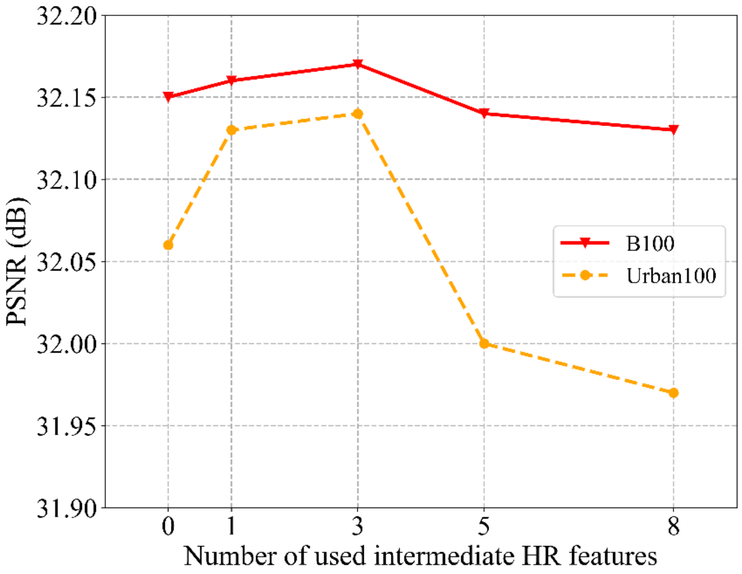
	CBPN-L	CBPN-H	CBPN
HR features		√	√
LR features	√		√
Set5	37.87	37.86	**37.90**
B100	32.15	32.13	**32.17**
Urban100	32.06	32.10	**32.14**

SR accuracy in terms of PSNR (dB) of our CBPN with or without using LR/HR features on three benchmark datasets for $\times 2$ SR.

Experiments

- Ablation study





The model gets best SR results when T=3.

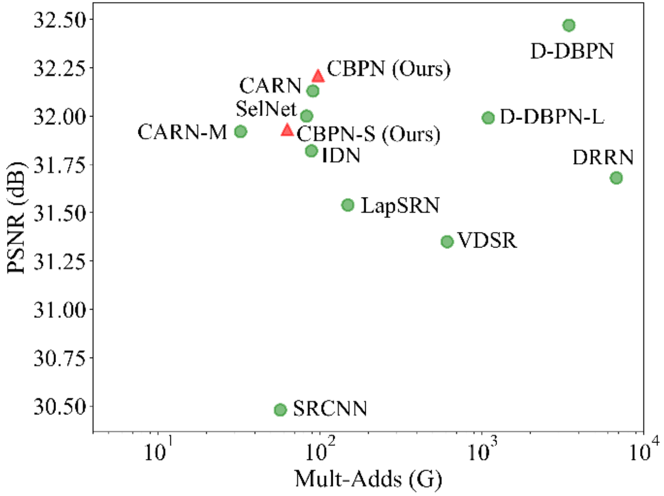
SR accuracy in terms of PSNR of our CBPN for $\times 2$ SR on B100 and Urban100 datasets w.r.t. the number of used intermediate HR residual features generated by the UD blocks

Experiments

- Trade-off between SR accuracy (PSNR) and the number of operations

Algorithm	Mult-Adds	Set5 PSNR/SSIM	Set14 PSNR/SSIM
D-DBPN-L [7]	1101.7G	31.99/0.893	28.52/0.778
CBPN	97.9G	32.21/0.894	28.63/0.781

Quantitative comparison results between our CBPN and D-DBPN-L for × 4 SR.



Experiments

- Quantitative results on X2 scale

Scale	Model	Params	Mult-Adds	Set5 PSNR/SSIM	Set14 PSNR/SSIM	B100 PSNR/SSIM	Urban100 PSNR/SSIM
	SRCNN [5]	57K	52.7G	36.66/0.9542	32.42/0.9063	31.36/0.8879	29.50/0.8946
	VDSR [10]	665K	612.6G	37.53/0.9587	33.03/0.9124	31.90/0.8960	30.76/0.9140
	LapSRN [12]	813K	29.9G	37.52/0.9590	33.08/0.9130	31.80/0.8950	30.41/0.9100
	DRRN [17]	297K	6,796.9G	37.74/0.9591	33.23/0.9136	32.05/0.8973	31.23/0.9188
	SelNet [3]	974K	225.7G	37.89/0.9598	33.61/0.9160	32.08/0.8984	-
	IDN [9]	553K	202.8G	37.83/0.9600	33.30/0.9148	32.08/0.8985	31.27/0.9196
×2	CARN [1]	1,592K	222.84G	37.76/0.9590	33.52/0.9166	32.09/0.8978	31.92/0.9256
	CARN-M [1]	412K	91.2G	37.53/0.9583	33.26/0.9141	31.92/0.8960	30.83/0.9233
	FALSR-A [4]	1,021K	234.7G	37.82/0.9595	33.55/0.9168	32.12/0.8987	31.93/0.9256
	FALSR-B [4]	326K	74.7G	37.61/0.9585	33.29/0.9143	31.97/0.8967	31.28/0.9191
	FALSR-C [4]	408K	93.7G	37.66/0.9586	33.26/0.9140	31.96/0.8965	31.24/0.9187
	CBPN (Ours)	1,036K	240.7G	37.90/0.9590	33.60/0.9171	32.17/0.8989	32.14/0.9279
	CBPN-S (Ours)	430K	101.5G	37.69/0.9583	33.36/0.9147	32.02/0.8972	31.55/0.9217

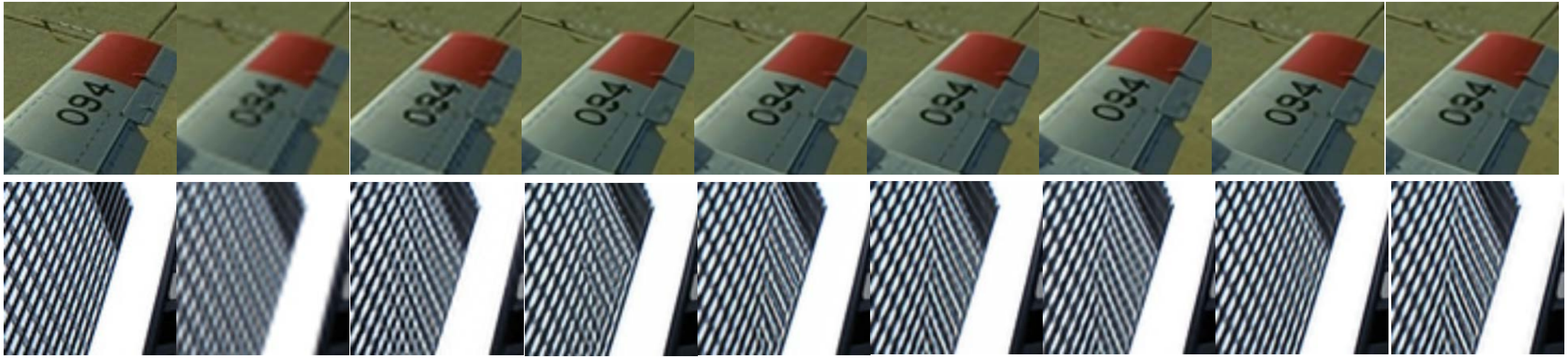
Experiments

- Quantitative results on X4 scale

Scale	Model	Params	Mult-Adds	Set5 PSNR/SSIM	Set14 PSNR/SSIM	B100 PSNR/SSIM	Urban100 PSNR/SSIM
×4	SRCNN [5]	57K	52.7G	30.48/0.8628	27.49/0.7503	26.90/0.7101	24.52/0.7221
	VDSR [10]	665K	612.6G	31.35/0.8838	28.01/0.7674	27.29/0.7251	25.18/0.7524
	DRCN [11]	1,774K	9,788.7G	31.53/0.8854	28.02/0.7670	27.23/0.7233	25.14/0.7510
	LapSRN [12]	813K	149.4G	31.54/0.8850	28.19/0.7720	27.32/0.7280	25.21/0.7560
	DRRN [17]	297K	6,796.9G	31.68/0.8888	28.21/0.7720	27.38/0.7284	25.44/0.7638
	SelNet [3]	1,417K	83.1G	32.00/0.8931	28.49/0.7783	27.44/0.7325	-
	IDN [9]	553K	89.0G	31.82/0.8903	28.25/0.7730	27.41/0.7297	25.41/0.7632
	SRDenseNet [21]	2,015K	389.9G	32.02/0.8934	28.50/0.7782	27.53/0.7337	26.05/0.7819
	CARN [1]	1,592K	90.9G	32.13/0.8937	28.60/0.7806	27.58/0.7349	26.07/0.7837
	CARN-M [1]	412K	32.5G	31.92/0.8903	28.42/0.7762	27.44/0.7304	25.63/0.7688
	CBPN (Ours)	1,197K	97.9G	32.21/0.8944	28.63/0.7813	27.58/0.7356	26.14/0.7869
	CBPN-S (Ours)	592K	63.1G	31.93/0.8908	28.50/0.7785	27.50/0.7324	25.85/0.7772

Experiments

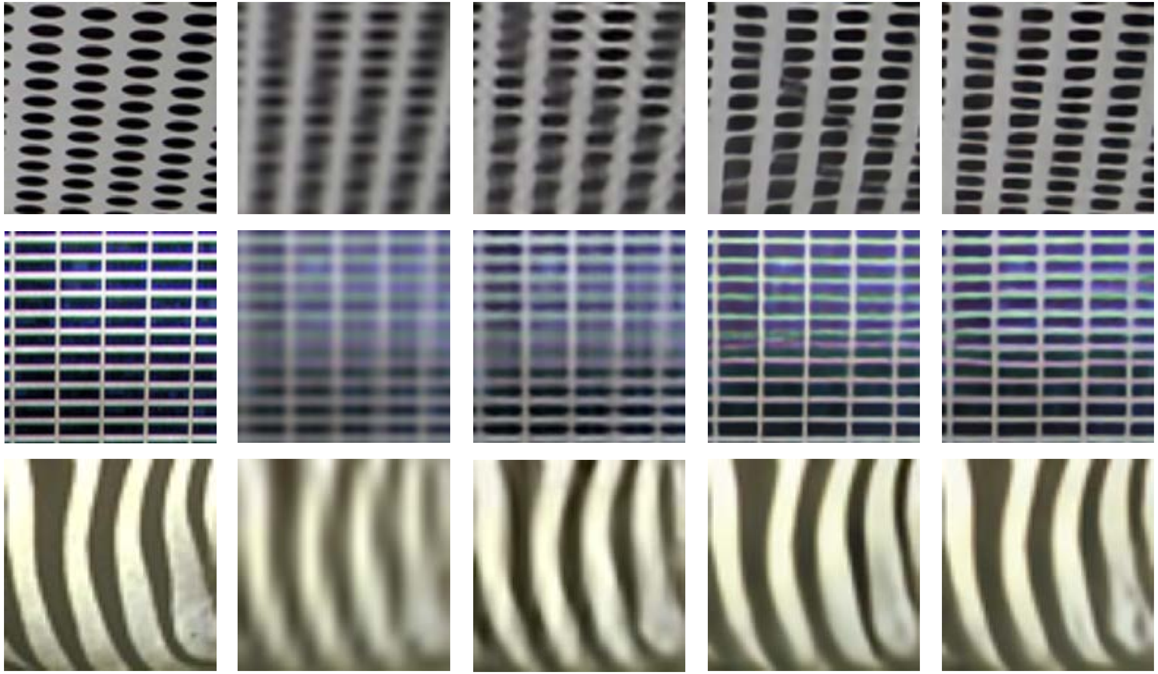
- Visual comparison on $\times 2$ scale



HR	Bicubic	SRCNN	FALSR-A	CARN	CARN-M	IDN	**CBPN**	**CBPN-S**

Experiments

- Visual comparison on $\times$ 4 scale



HR Bicubic SRCNN CARN **CBPN**

Conclusion

- We propose **compact back-projection networks (CBPN)** for efficient single image super-resolution.
- **UD-block** ensures the efficiency of our model and **hybrid residual features** information are used in the final reconstruction.
- We use **global and local residual connection** to promote our network learning residual images between HR images and interpolated images.
- **Compression layers** are employed to fusion features and reduce the number of parameters and operations of our network.

Q & A

Thanks!