Thirty-Sixth AAAI Conference on Artificial Intelligence 2022

Feature Distillation Interaction Weighting Network for Lightweight Image Super-Resolution

Guangwei Gao[1†], Wenjie Li[1†], Juncheng Li[2*], Fei Wu[1], Huimin Lu[3], Yi Yu[4]

[1]Nanjing University of Posts and Telecommunications [2] The Chinese University of Hong Kong
[3] Kyushu Institute of Technology [4] National Institute of Informatics

Reporter: Wenjie Li

https://github.com/IVIPLab/FDIWN

Outline

- Background & Related Works
- Motivation
- Feature Distillation Interaction Weighting Network
- Experiments & Discussion
- Summary

Outline

- Background & Related Works
- Motivation
- Feature Distillation Interaction Weighting Network
- Experiments & Discussion
- Summary

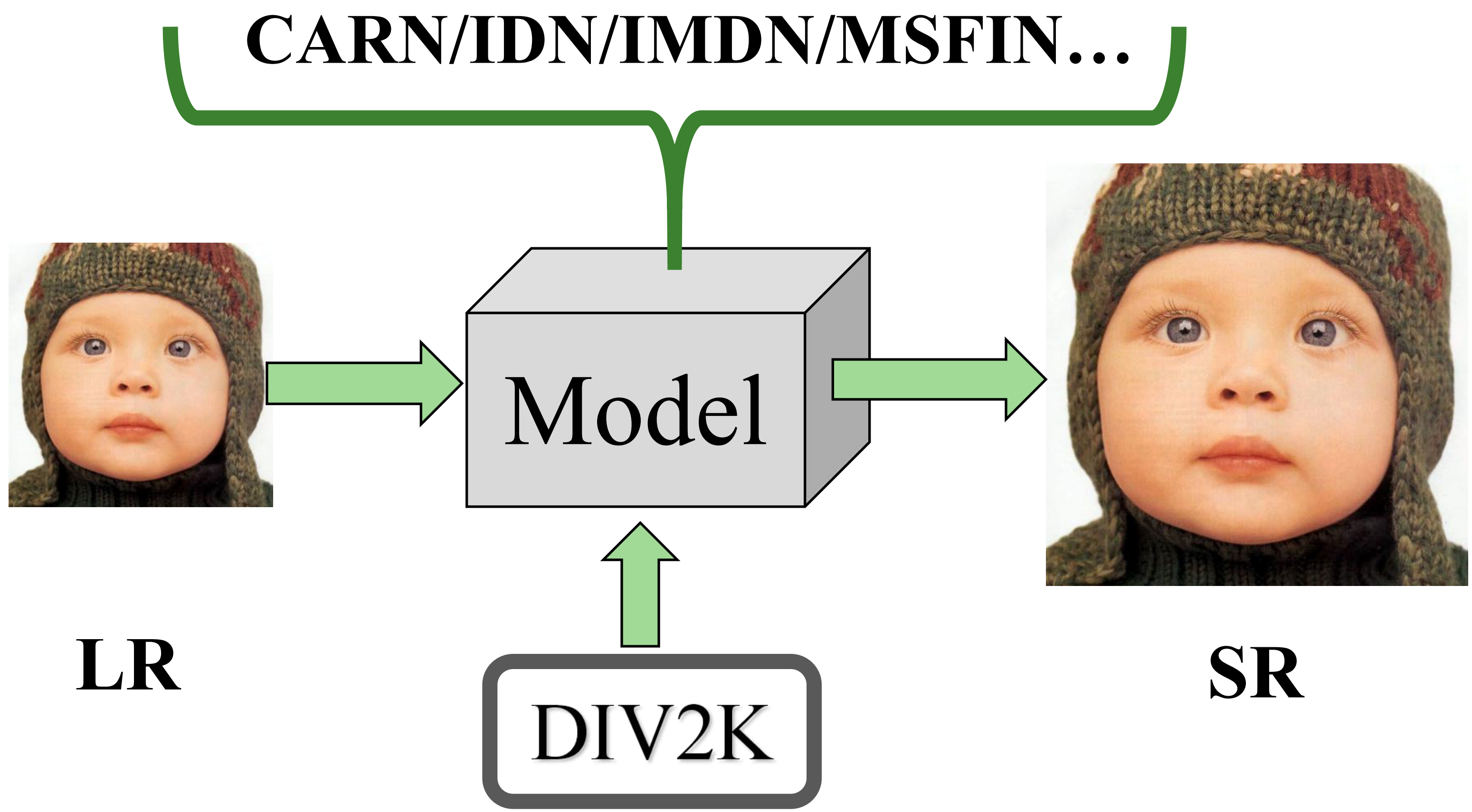
Background & Related Works

- Image Super-resolution
  - The purpose of single-image super-resolution (SISR) is **to reconstruct a high-resolution (HR) image from its degraded low-resolution (LR) counterpart**.

- Lightweight Image Super-resolution
  - **Lightweight** SR models are widely concerned for saving memory resources and computing resources.
  - In the case of **fewer parameters and computation,** a better performance is obtained.

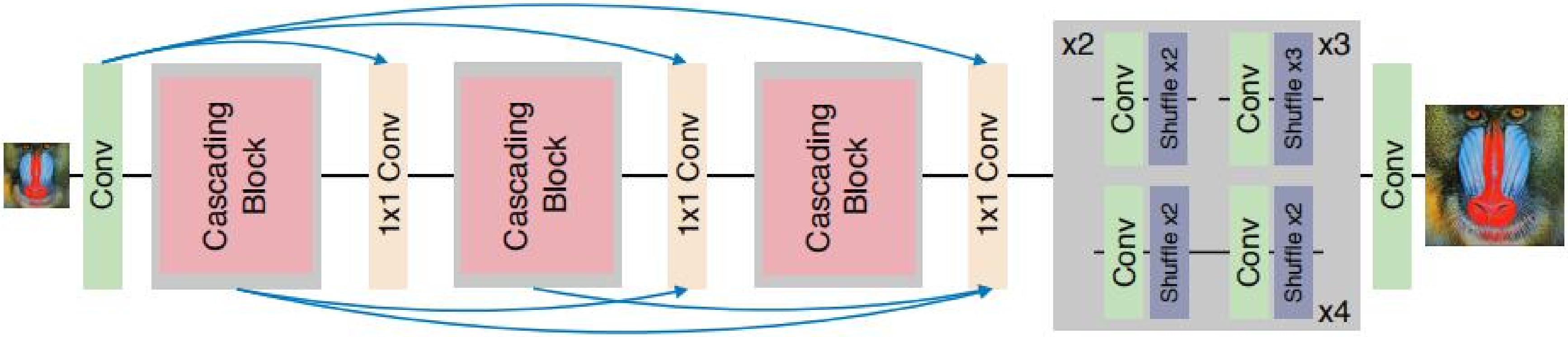
Background & Related Works

- How to reconstruct SR images?



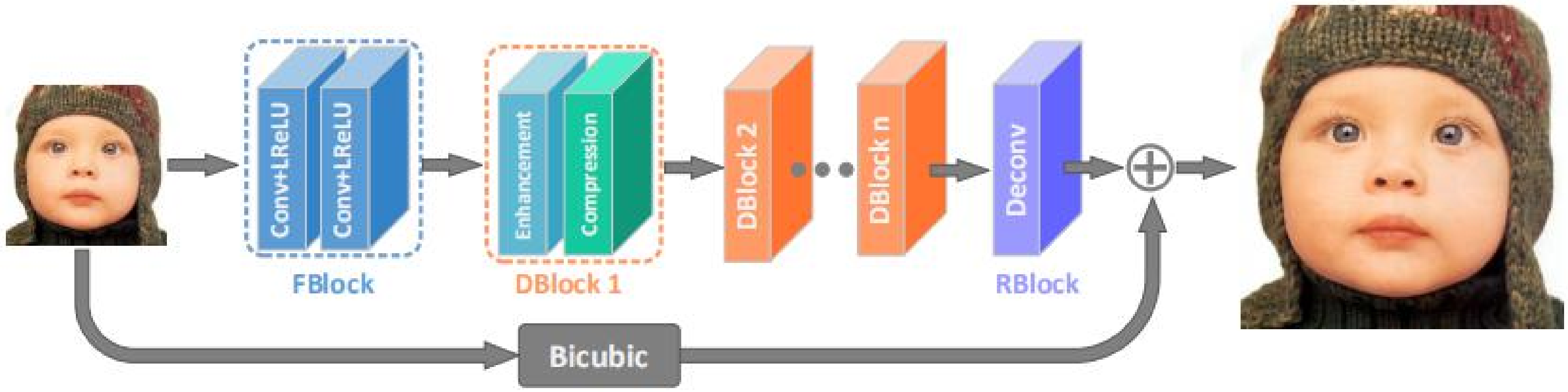
Background & Related Works

- Ahn's CARN



N. Ahn, et al. Fast, accurate, and lightweight super-resolution with cascading residual network, in *ECCV,* 2018, pp. 252–268.

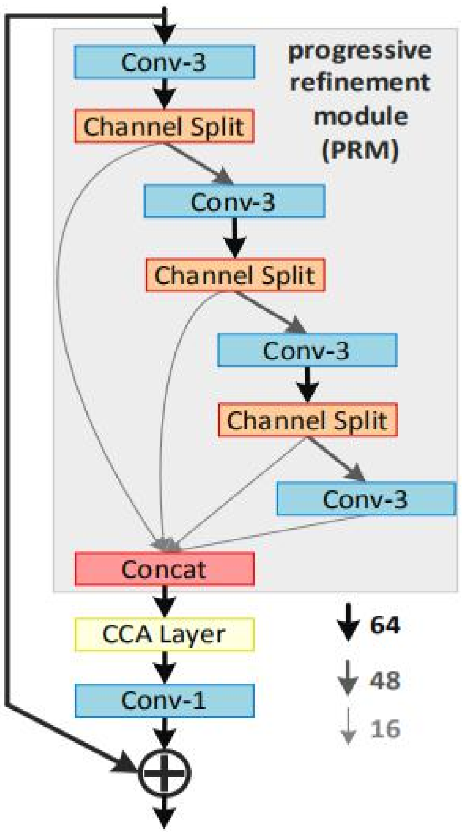
- Hui's IDN



Z. Hui, et al. Fast and accurate single image super-resolution via information distillation network, in *CVPR*, 2018, pp. 723–731.

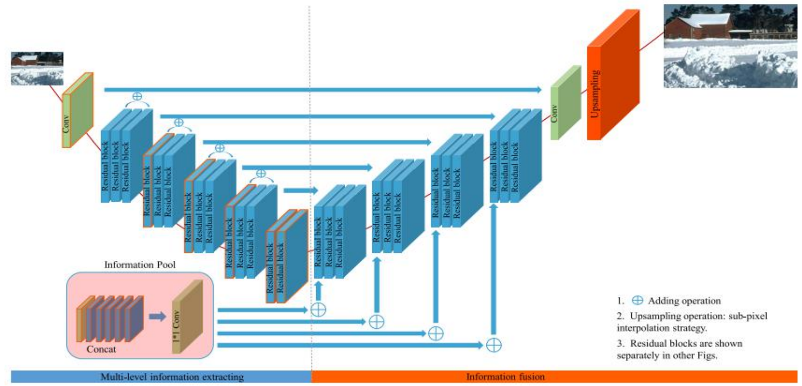
Background & Related Works

- Hui's



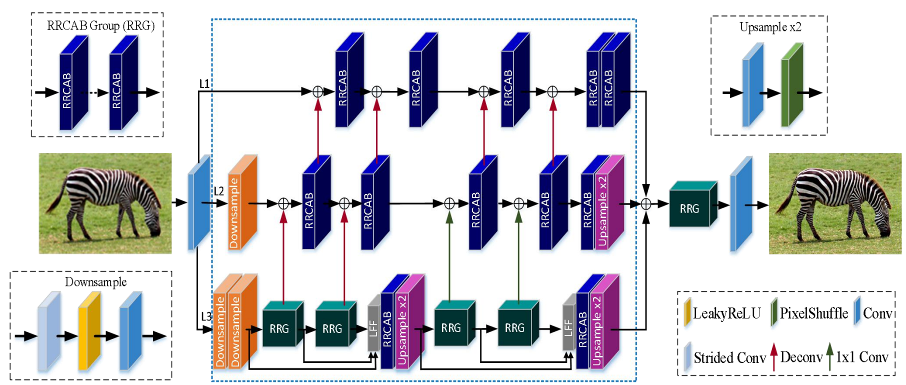
Z. Hui, et al. Lightweight image super-resolution with information multi-distillation network, in *ACM MM*, 2019, pp. 2024–2032.

- Li's s-LWSR



B. Li, et al. s-lwsr: Super lightweight super-resolution network, *IEEE TIP*, vol. 29, pp. 8368–8380, Aug. 2020.

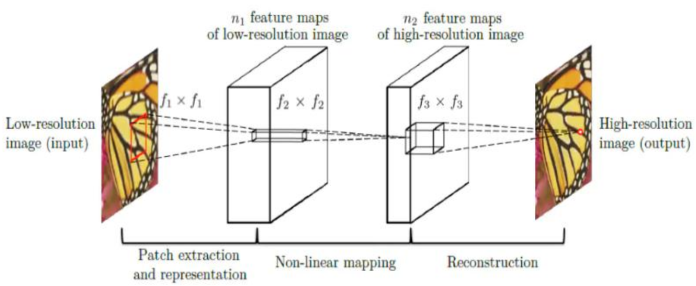
- Wang's MSFIN



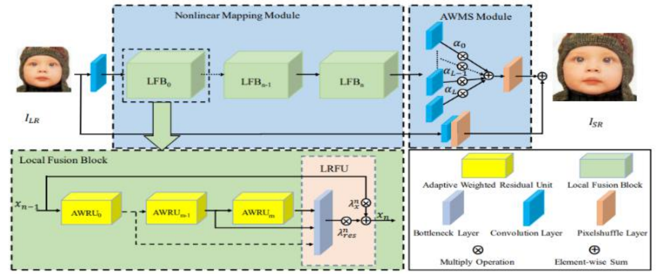
B. Li, et al. s-lwsr: Super lightweight super-resolution network, *IEEE TIP*, vol. 29, pp. 8368–8380, Aug. 2020.

Background & Related Works

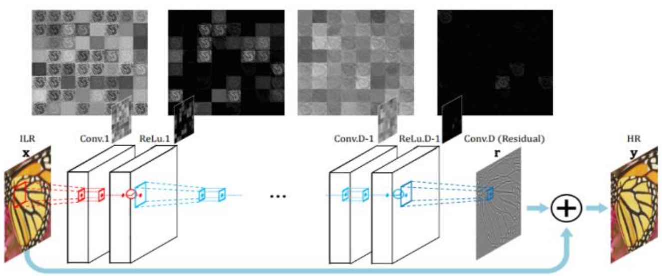
- Other Lightweight Methods



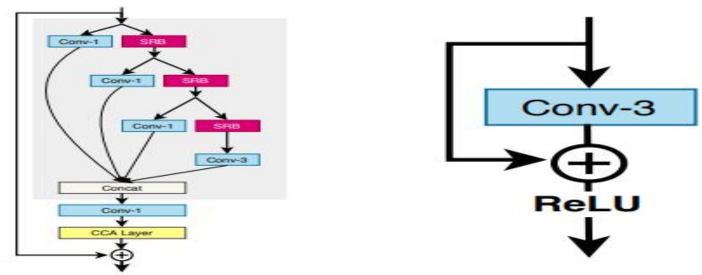
SRCNN



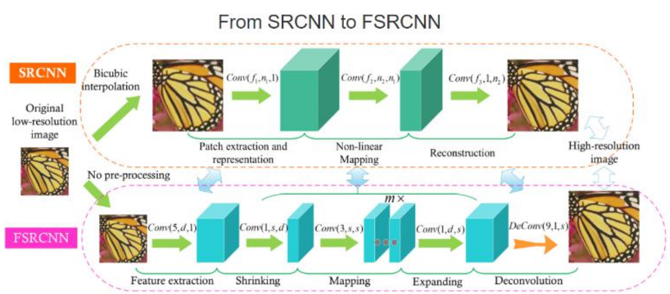
AWSRN



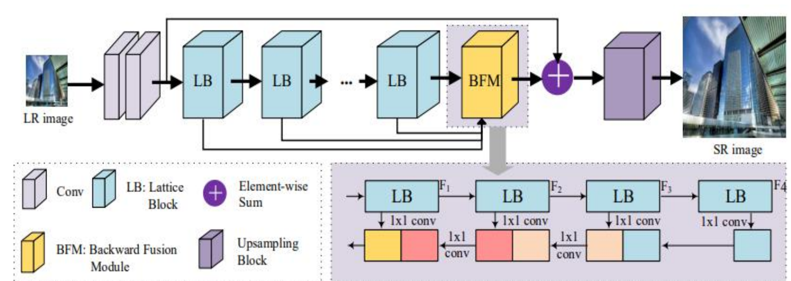
VDSR



RFDN



FSRCNN



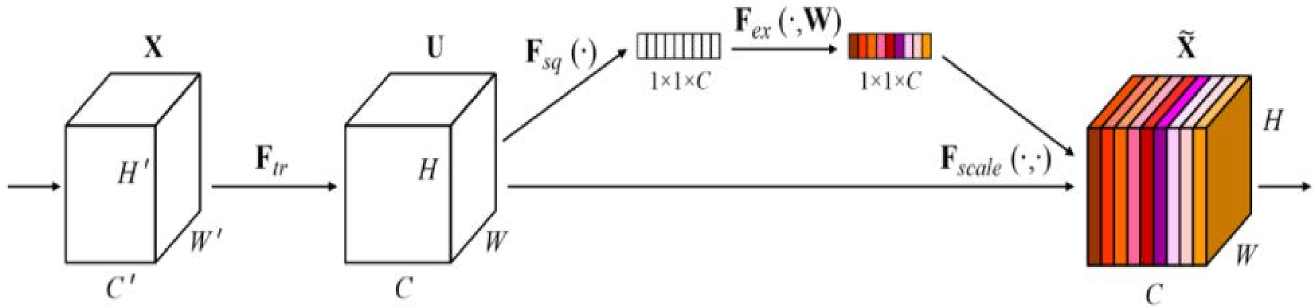
LatticeNet

Outline

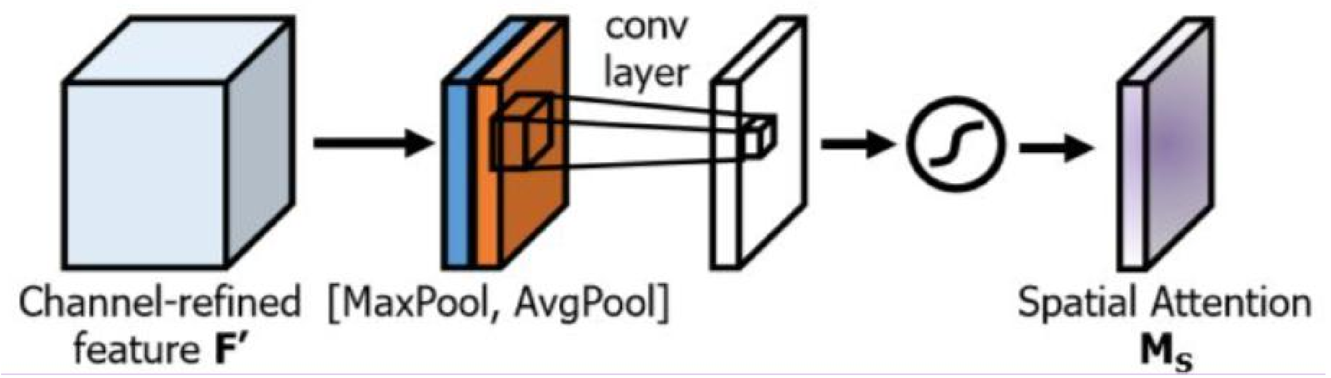
- Background & Related Works
- Motivation
- Feature Distillation Interaction Weighting Network
- Experiments & Discussion
- Summary

Motivation

- Attention Mechanism



Channel Attention

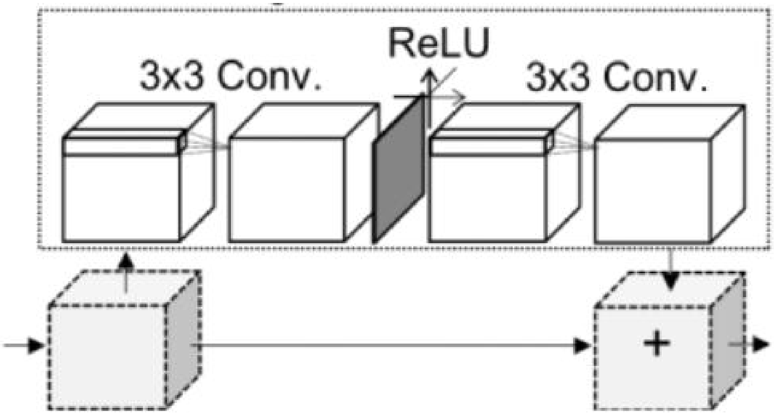


Spatial Attention

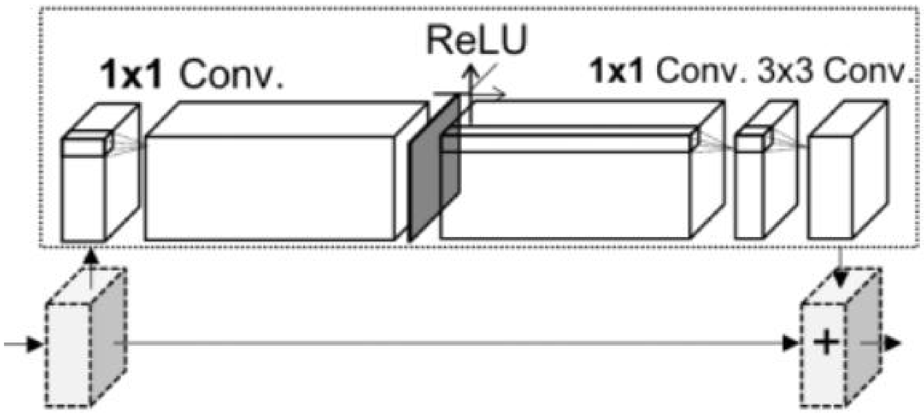
In our proposed model, We believe that we should take advantage of both channel attention and spatial attention to obtain better SR performance.

Motivation

- Stronger feature extraction capabilities



Original ResBlock



WDSR

Since WDSR has a broad activation mechanism, its information extraction ability will be better. In order for the model to focus on the important information, why not make the SA attention focus on each WDSR unit and adjust the output adaptively ?

Motivation

AAAI-22

Target

- We aim to explore a lightweight and efficient SISR model.

Contributions

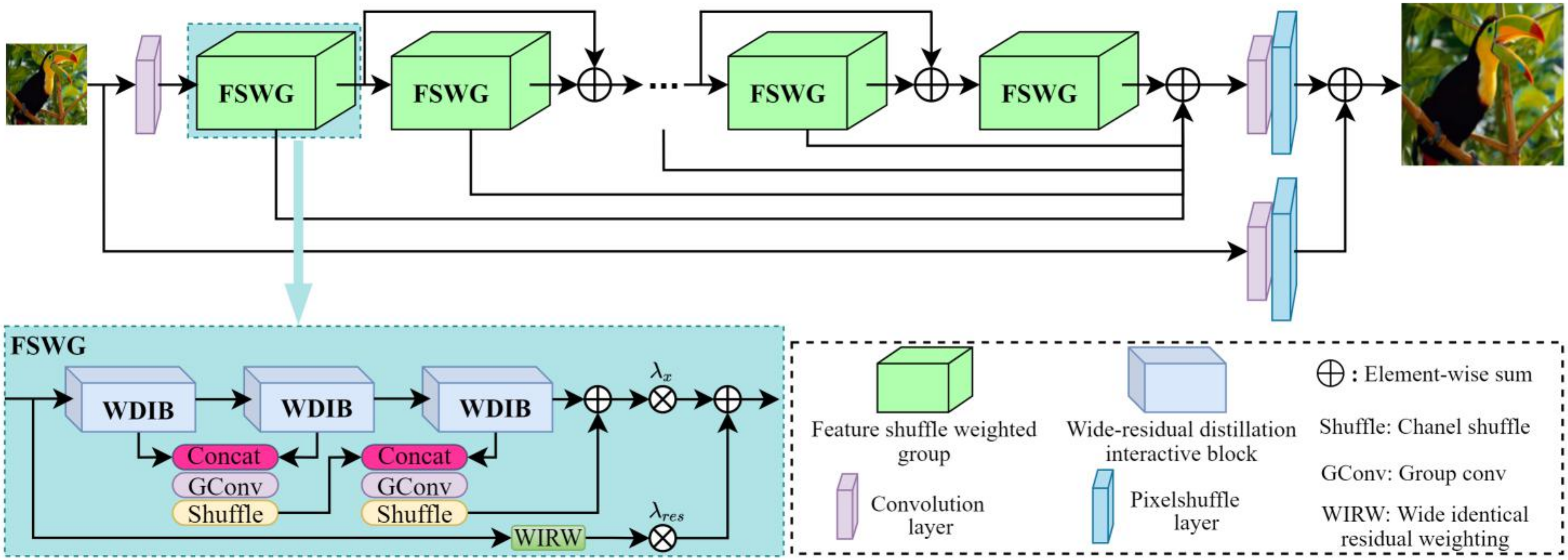
- We propose a wide-residual attention weighting unit for lightweight SISR, including Wide Identical Residual Weighting (WIRW) unit and Wide Convolutional Residual Weighting (WCRW) unit, which has stronger feature distillation capabilities than ordinary residual blocks.
- We propose a novel Self-Calibration Fusion (SCF) module to replace the traditional concatenate operation for efficient feature interaction and fusion, which can aggregate more representative features and self-calibrate the input and output features.
- We propose a Wide-Residual Distillation Connection (WRDC) framework, which connects the coarse and distilled fine features within the module and allows features from different scales to interact with each other.
- We design a Feature Shuffle Weighted Group (FSWG) for pairwise feature fusion, which consists of a series of interactional WDIBs.

Outline

- Background & Related Works
- Motivation
- Feature Distillation Interaction Weighting Network
- Experiments & Discussion
- Summary

FDIWN

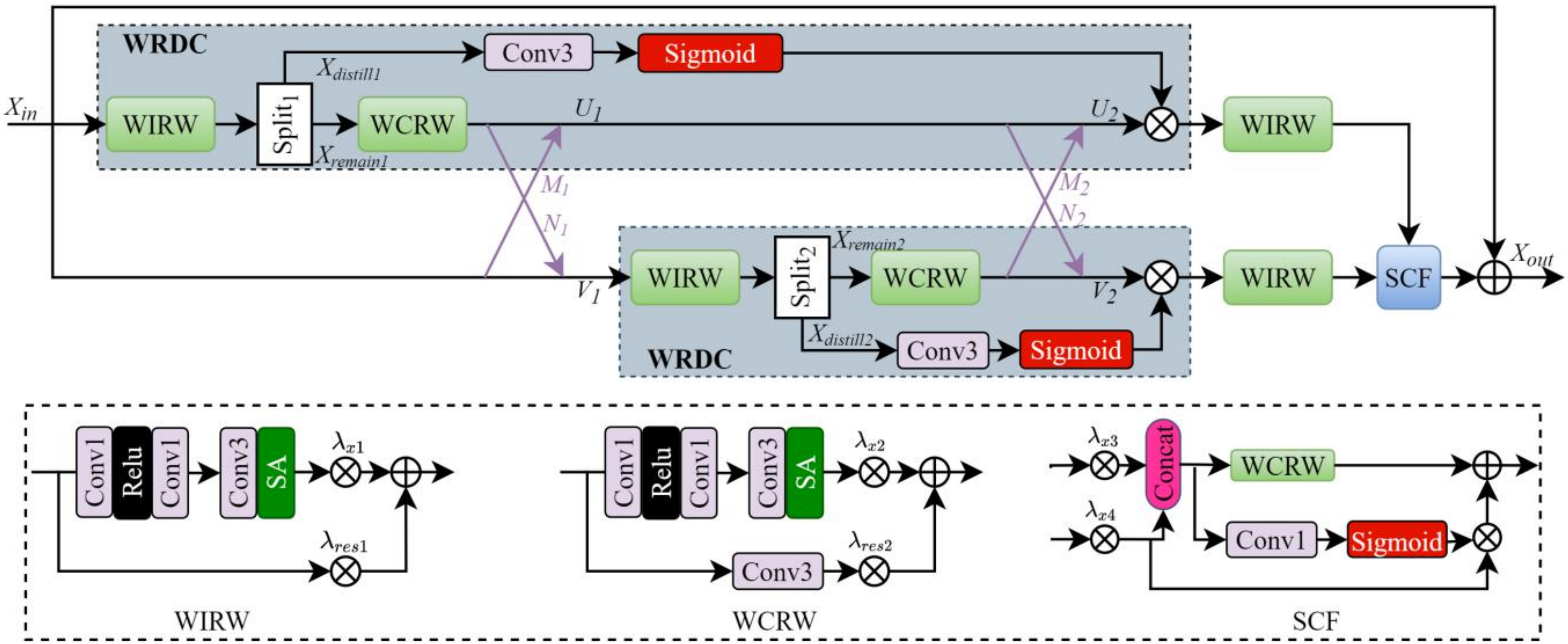
- Network Architecture



Loss function: $$\hat{\theta} = arg \min_{\theta} \frac{1}{N} \sum_{i=1}^{N} \left\| F_{\theta}(I_{LR}^{i}) - I_{HR}^{i} \right\|_1$$

FDIWN-WDIB

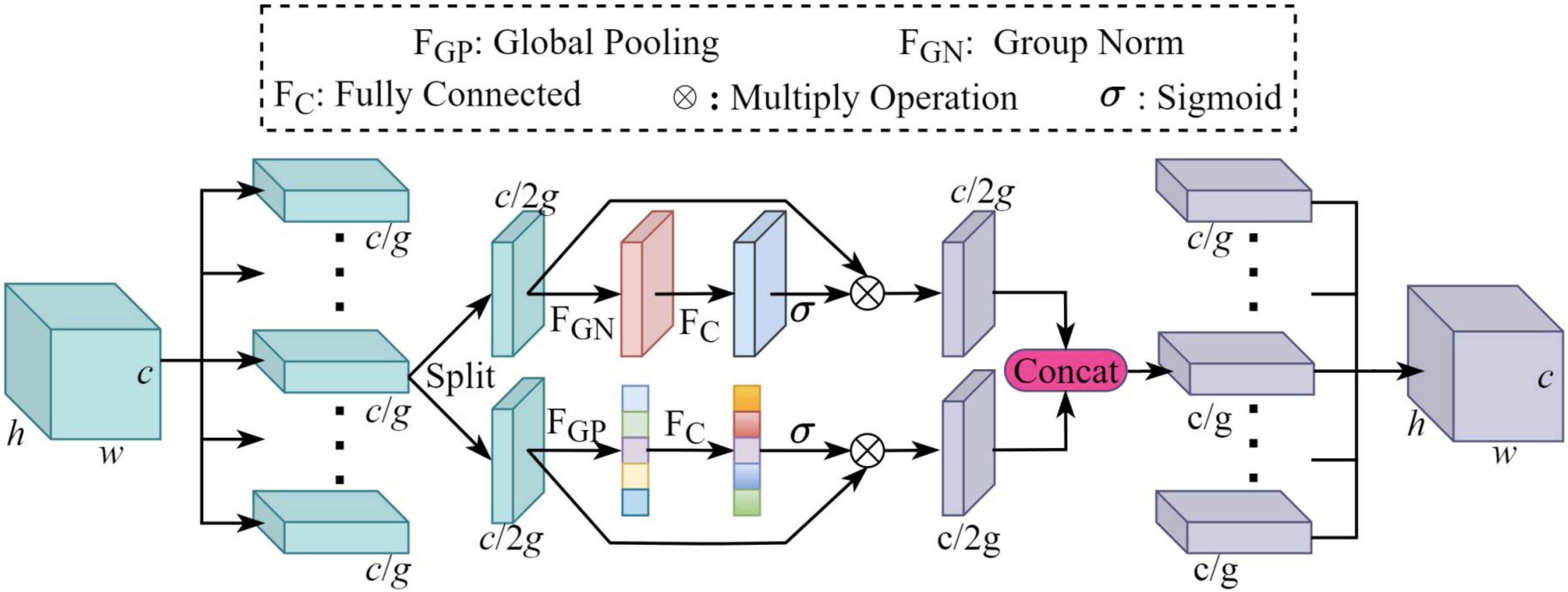
- Wide-residual Distillation Interaction Block



- The Mi and Ni is the combination coefficient learning.
- The SA is the Shuffle Attention.

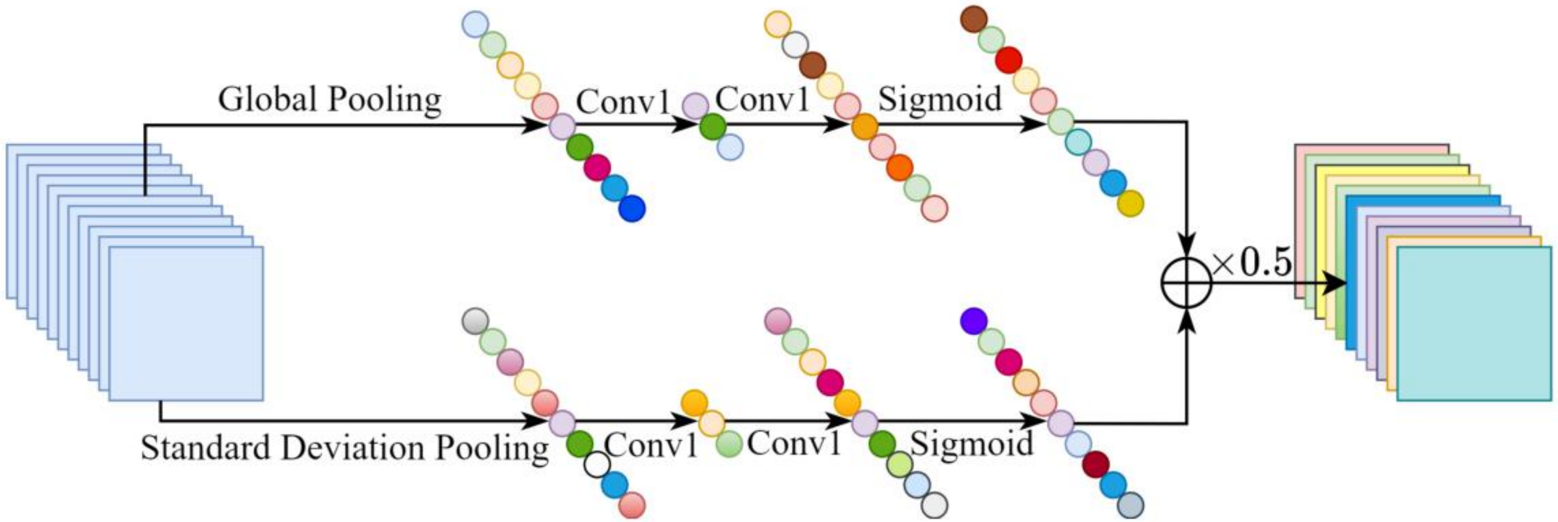
FDIWN-SA

- Shuffle Attention Mechanism



FDIWN

- Combination Coefficient Learning



$$H_{GC}(x_c) = \sqrt{\frac{1}{HW}\sum_{(i,j)\in x_c}(x_c^{i,j} - \frac{1}{HW}\sum_{(i,j)\in x_c} x_c^{i,j})^2}$$

Outline

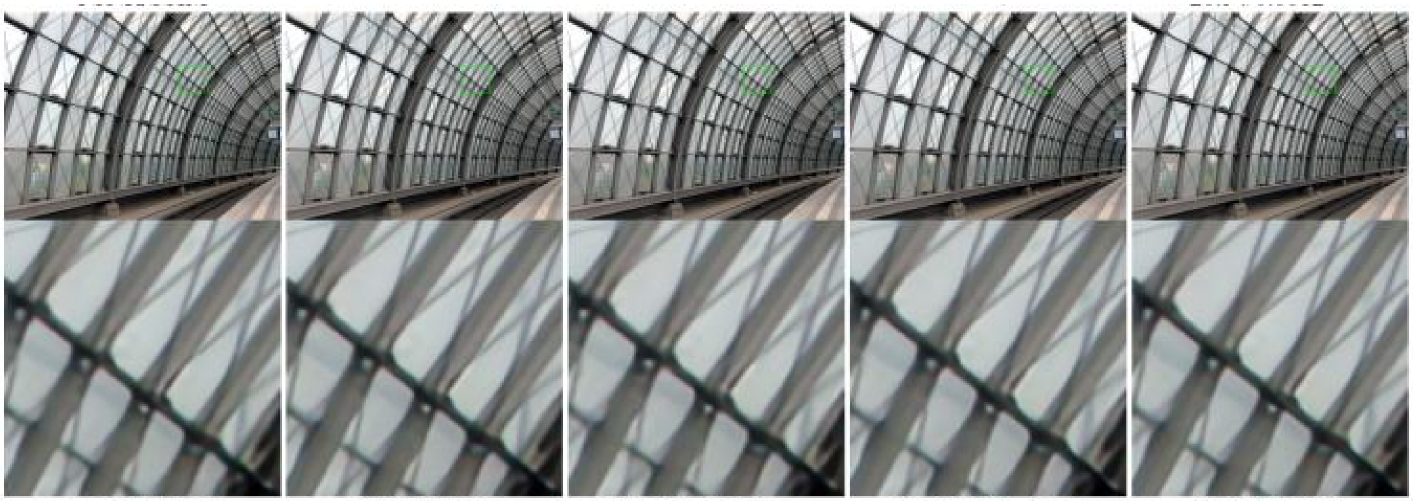
- Background & Related Works
- Motivation
- Feature Distillation Interaction Weighting Network
- Experiments & Discussion
- Summary

Experiments & Discussion

- Quantitative comparisons

Algorithm	Scale	Params	Multi-adds	Set5 PSNR	Set5 SSIM	Set14 PSNR	Set14 SSIM	BSDS100 PSNR	BSDS100 SSIM	Urban100 PSNR	Urban100 SSIM
SRCNN (Dong et al. 2015)		57K	52.7G	32.75	0.9090	29.30	0.8215	28.41	0.7863	26.24	0.7889
FSRCNN (Dong, Loy, and Tang 2016)		12K	5.0G	33.16	0.9140	29.43	0.8242	28.53	0.7910	26.43	0.8080
VDSR (Kim, Lee, and Lee 2016a)		665K	612.6G	33.67	0.9210	29.78	0.8320	28.83	0.7990	27.14	0.8290
DRCN (Kim, Lee, and Lee 2016b)		1774K	17974.3G	33.82	0.9226	29.76	0.8311	28.80	0.7963	27.15	0.8276
IDN (Hui, Wang, and Gao 2018)		590K	105.6G	34.11	0.9253	29.99	0.8354	28.95	0.8013	27.42	0.8359
CARN-M (Ahn, Kang, and Sohn 2018)		412K	46.1G	33.99	0.9236	30.08	0.8367	28.91	0.8000	27.55	0.8385
CARN (Ahn, Kang, and Sohn 2018)		1592K	118.8G	34.29	0.9255	30.29	0.8407	29.06	0.8034	28.06	0.8493
IMDN (Hui et al. 2019)	×3	703K	71.5G	34.36	0.9270	30.32	0.8417	29.09	0.8046	28.17	0.8519
AWSRN-M (Wang, Li, and Shi 2019)		1143K	116.6G	34.42	0.9275	30.32	0.8419	29.13	0.8059	28.26	0.8545
MADNet (Lan et al. 2020)		930K	88.4G	34.16	0.9253	30.21	0.8398	28.98	0.8023	27.77	0.8439
RFDN (Liu, Tang, and Wu 2020)		541K	55.4G	34.41	0.9273	30.34	0.8420	29.09	0.8050	28.21	0.8525
MAFFSRN (Muqeet et al. 2020)		418K	34.2G	34.32	0.9269	30.35	0.8429	29.09	0.8052	28.13	0.8521
LAPAR-A (Li et al. 2021)		594K	114G	34.36	0.9267	30.34	0.8421	29.11	0.8054	28.15	0.8523
FDIWN-M(Ours)		446K	35.9 G	34.46	0.9274	30.35	0.8423	29.10	0.8051	28.16	0.8528
FDIWN(Ours)		645K	51.5G	34.52	0.9281	30.42	0.8438	29.14	0.8065	28.36	0.8567

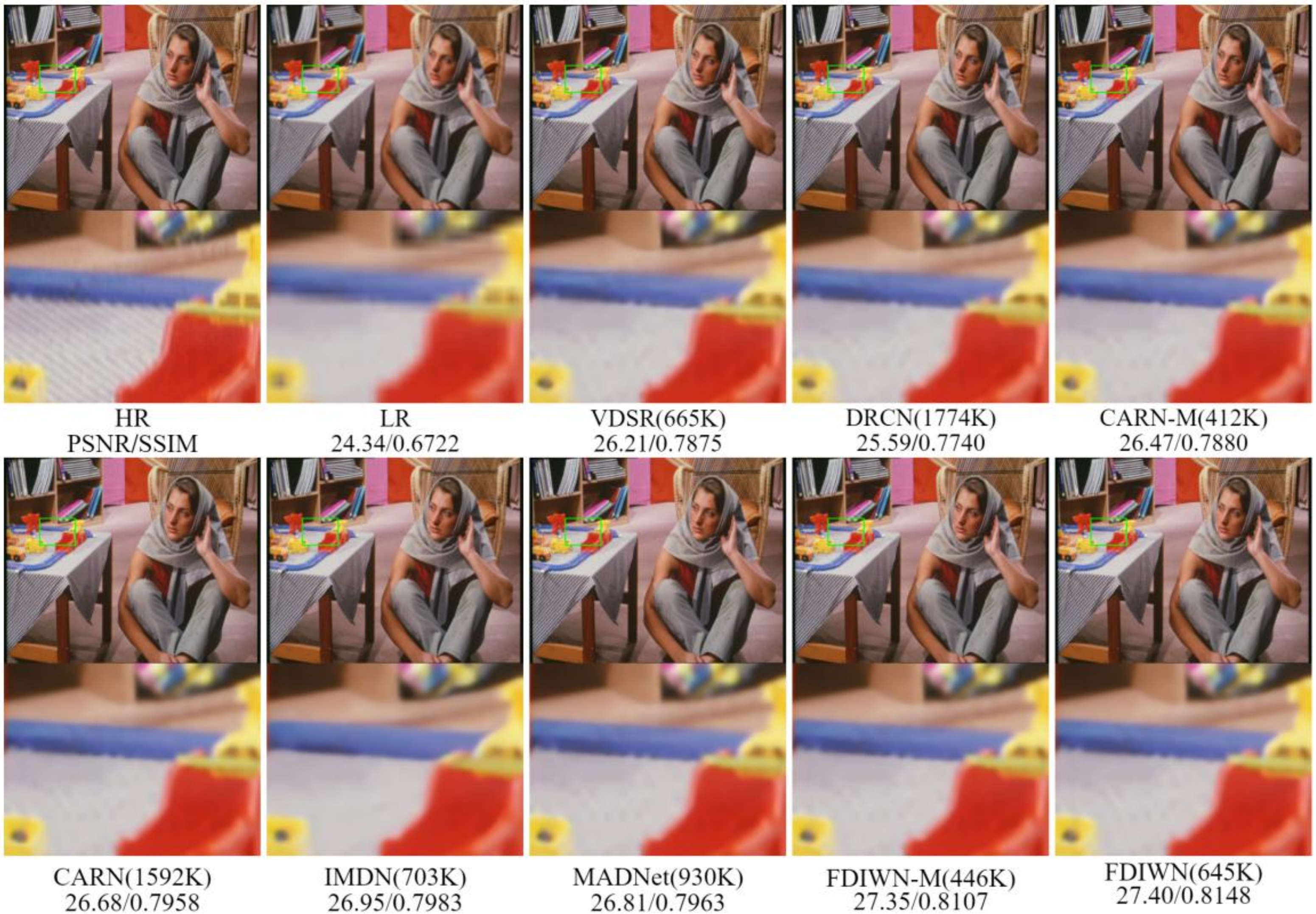
- Visual results of FDIWN with other SR methods (x3)



CARN(1592K) 29.11/0.8917 | IMDN(703K) 29.08/0.8922 | MADNet(930K) 28.70/0.8840 | FDIWN-M(446K) 29.09/0.8920 | FDIWN(645K) 29.54/0.8998

Experiments & Discussion

- Visual results of FDIWN with other SR methods (x3)



Experiments & Discussion

- Quantitative comparisons

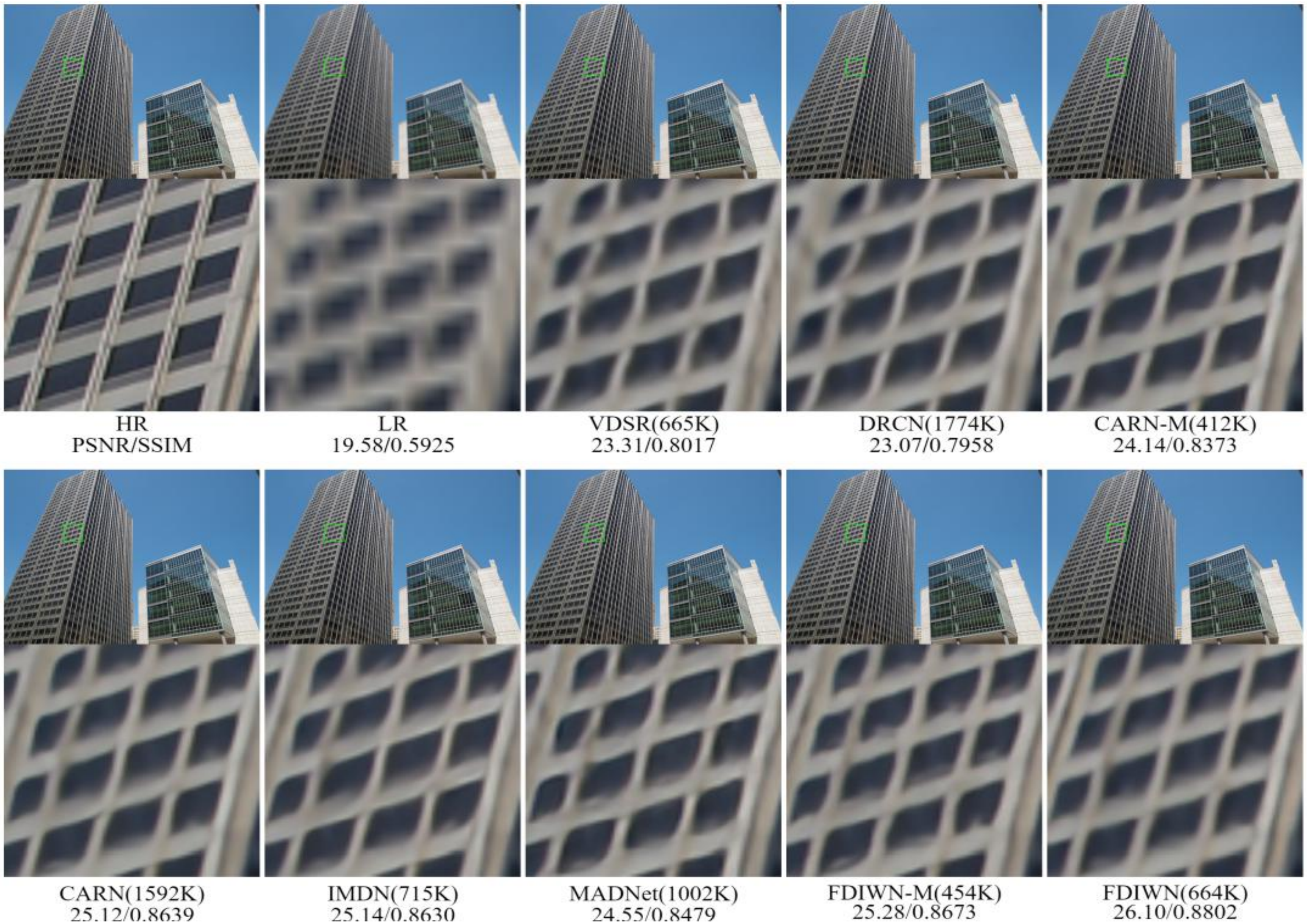
SRCNN (Dong et al. 2015)		57K	52.7G	30.48	0.8628	27.49	0.7503	26.90	0.7101	24.52	0.7221
FSRCNN (Dong, Loy, and Tang 2016)		12K	4.6G	30.71	0.8657	27.59	0.7535	26.98	0.7150	24.62	0.7280
VDSR (Kim, Lee, and Lee 2016a)		665K	612.6G	31.35	0.8838	28.01	0.7674	27.29	0.7251	25.18	0.7524
DRCN (Kim, Lee, and Lee 2016b)		1774K	17974.3G	31.53	0.8854	28.02	0.7670	27.23	0.7233	25.14	0.7510
LapSRN (Lai et al. 2017)		813K	149.4G	31.54	0.8850	28.19	0.7720	27.32	0.7280	25.21	0.7560
IDN (Hui, Wang, and Gao 2018)		590K	81.9G	31.82	0.8903	28.25	0.7730	27.41	0.7297	25.41	0.7632
CARN-M (Ahn, Kang, and Sohn 2018)		412K	32.5G	31.92	0.8903	28.42	0.7762	27.44	0.7304	25.62	0.7694
CARN (Ahn, Kang, and Sohn 2018)		1592K	90.9G	32.13	0.8937	28.60	0.7806	27.58	0.7349	26.07	0.7837
IMDN (Hui et al. 2019)	$\times 4$	715K	40.9G	32.21	0.8948	28.58	0.7811	27.56	0.7353	26.04	0.7838
AWSRN-M (Wang, Li, and Shi 2019)		1254K	72.0G	32.21	0.8954	28.65	0.7832	27.60	0.7368	26.15	0.7884
MADNet (Lan et al. 2020)		1002K	54.1G	31.95	0.8917	28.44	0.7780	27.47	0.7327	25.76	0.7746
RFDN (Liu, Tang, and Wu 2020)		550K	31.6G	32.24	0.8952	28.61	0.7819	27.57	0.7360	26.11	0.7858
MAFFSRN (Muqeet et al. 2020)		441K	19.3G	32.18	0.8948	28.58	0.7812	27.57	0.7361	26.04	0.7848
ECBSR (Zhang, Zeng, and Zhang 2021)		603K	34.73G	31.92	0.8946	28.34	0.7817	27.48	0.7393	25.81	0.7773
LAPAR-A (Li et al. 2021)		659K	94G	32.15	0.8944	28.61	0.7818	27.61	0.7366	26.14	0.7871
FDIWN-M(Ours)		454K	19.6G	32.17	0.8941	28.55	0.7806	27.58	0.7364	26.02	0.7844
FDIWN(Ours)		664K	28.4G	32.23	0.8955	28.66	0.7829	27.62	0.7380	26.28	0.7919

- Visual results of FDIWN with other SR methods (x4)

CARN(1592K) 33.03/0.9447 | IMDN(715K) 32.77/0.7983 | MADNet(1002K) 32.35/0.9406 | FDIWN-M(454K) 32.86/0.9440 | FDIWN(664K) 33.07/0.9452

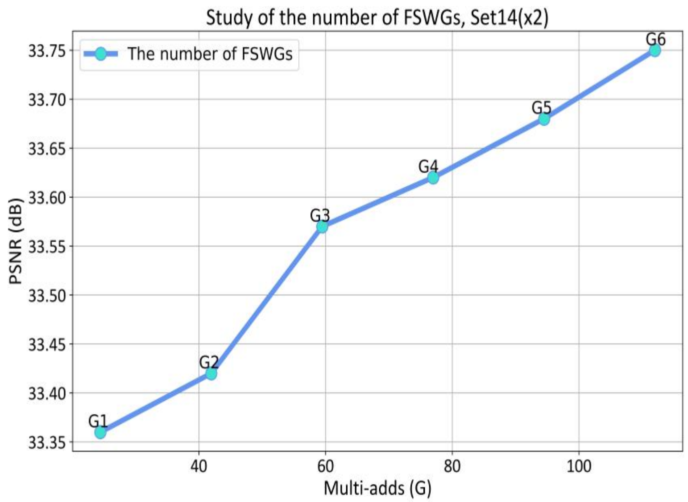
Experiments & Discussion

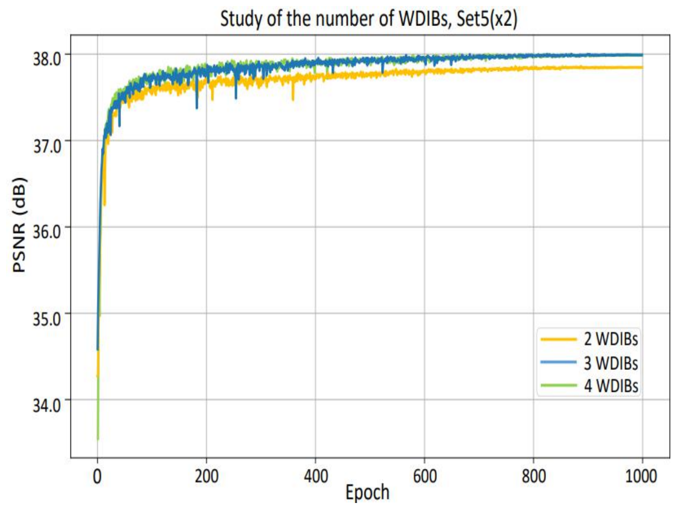
- Visual results of FDIWN with other SR methods (x4)



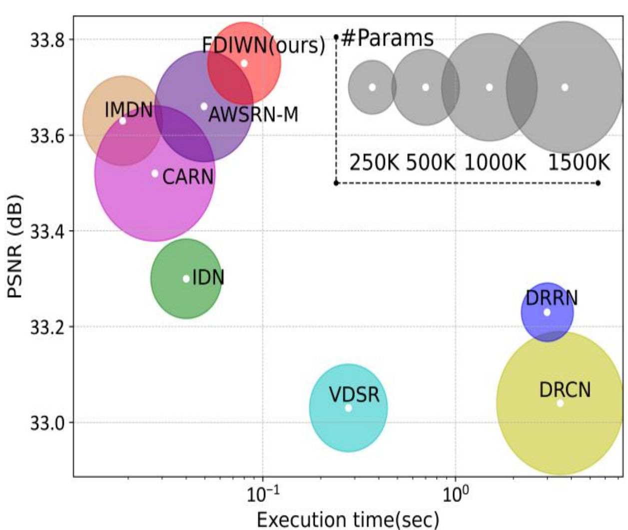
Experiments & Discussion

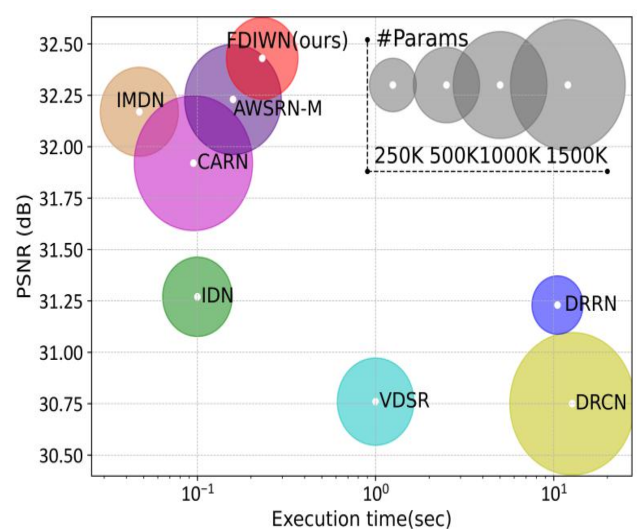
- Efficiency trade-off





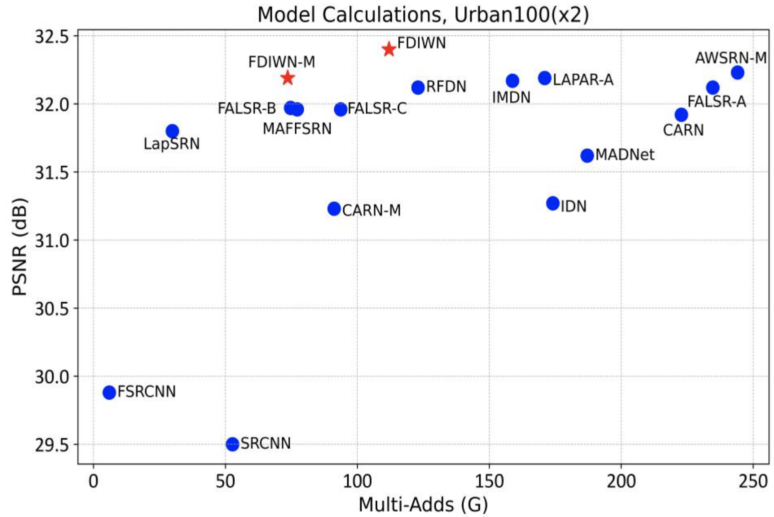
- Model complexity analysis



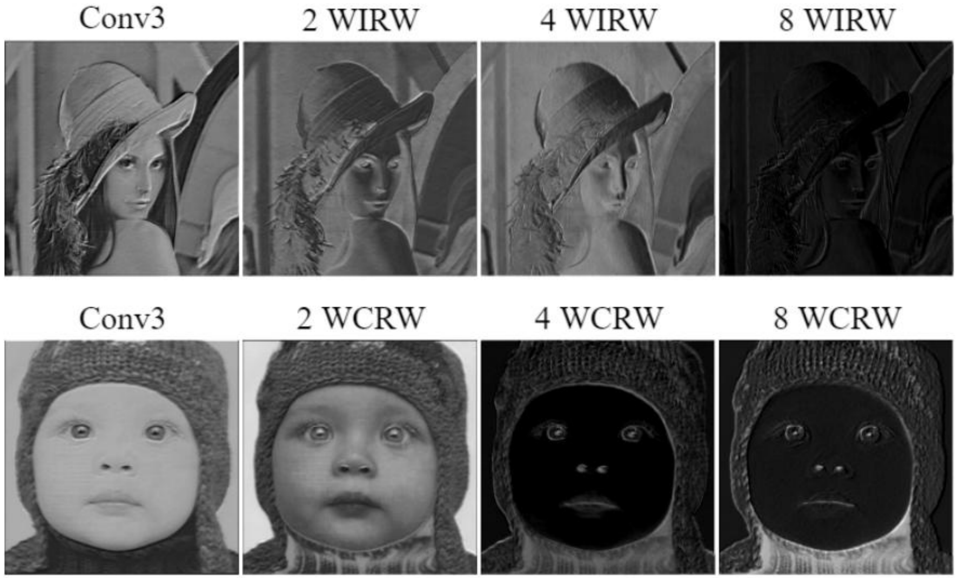


Experiments & Discussion

- Investigations of the model size and performance



- Feature visualization of different module



Experiments & Discussion

- The combination structure of WDIB

Method	BI	WIRW	Params	Multi-adds	PSNR	SSIM
Baseline	✗	✗	215K	22.0G	37.81	0.9598
FDIWN	✓	✗	225K	24.4G	37.85	**0.9600**
FDIWN	✓	✓	230K	24.4G	**37.88**	**0.9600**

- The effectiveness of WIRW and WCRW

Case	Method	Channels	Params	Multi-adds	PSNR	SSIM
1	Baseline	24	152K	23.2G	37.70	0.9594
2	FDIWN	48	96K	9.7G	37.64	0.9592
3	FDIWN	120	131K	9.7G	**37.72**	**0.9596**

- The effectiveness of WRDC and SCF

Method	WR	DC	SCF	Params	Multi-adds	PSNR	SSIM
Baseline1	✗	✗	✗	59K	3.3G	37.52	0.9587
Baseline2	✓	✗	✗	59K	4.9G	37.53	0.9589
FDIWN	✗	✗	✓	89K	6.5G	37.58	0.9591
FDIWN	✓	✓	✗	65K	6.5G	37.59	0.9590
FDIWN	✓	✓	✓	96K	9.7G	**37.64**	**0.9592**

Outline

- Background & Related Works
- Motivation
- Feature Distillation Interaction Weighting Network
- Experiments & Discussion
- Summary

Summary

- The specially designed wide-residual weighting units (including WIRW and WCRW) have a stronger ability to distill useful features than ordinary residual blocks.

- The well designed wide-residual units based WRDC module and SCF module can flexibly aggregate and distill more representative features.

- The experiment show that our proposed FDIWN achieved a good balance between model size, performance, and computational cost.

Thanks !