MemSR: Training Memory-efficient Lightweight Model for Image Super-Resolution

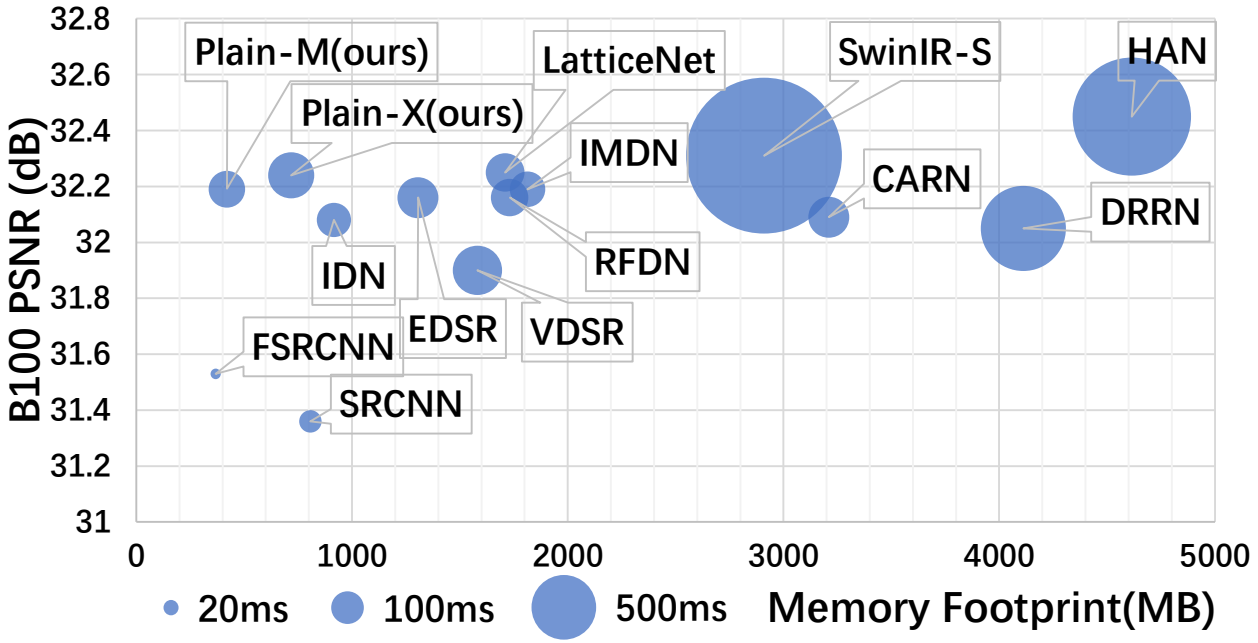
Kailu Wu, Chung-Kuei Lee, and Kaisheng Ma

International Conference on Machine Learning (ICML)

Spotlight Presentation

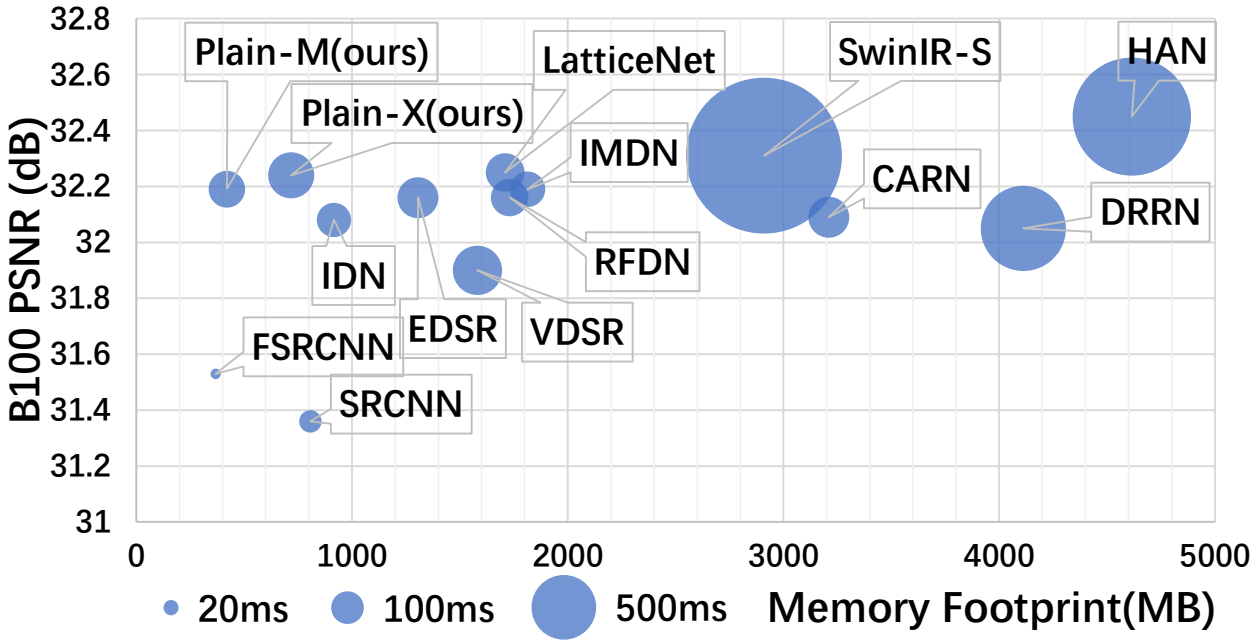
27 June 2022

Memory Problem



The maximum memory footprint and the average inference time for upscaling $2\times$ on LR image of size 960×540

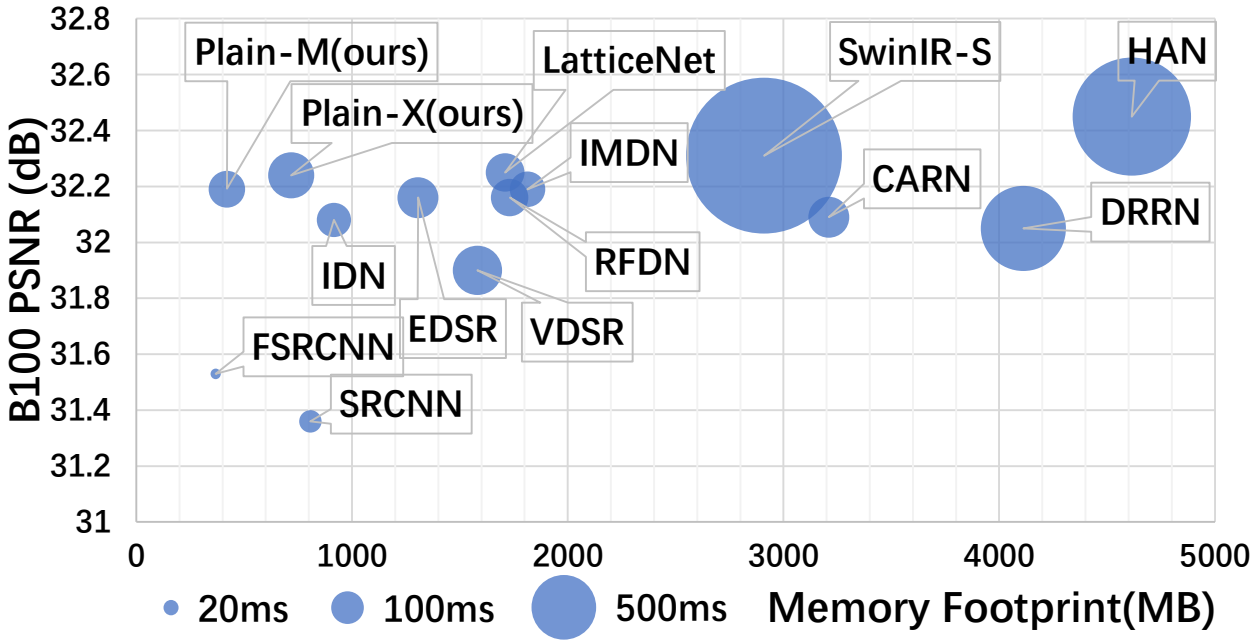
Memory Problem



The maximum memory footprint and the average inference time for upscaling $2\times$ on LR image of size 960×540

- Typical resolution on smart phone camera: $96MP$

Memory Problem



The maximum memory footprint and the average inference time for upscaling $2\times$ on LR image of size 960×540

- Typical resolution on smart phone camera: $96MP$
- The memory limitation for super-resolution algorithms becomes non-negligible on edge devices.

Solution?

- Plain model?

Solution?

- Plain model?
- How to train a strong plain model in super-resolution?

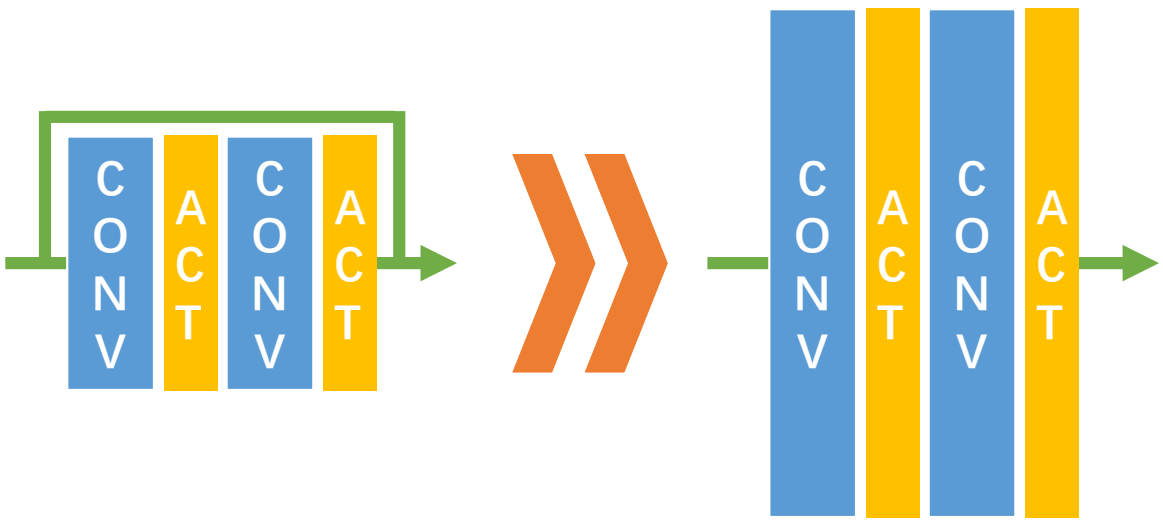
Solution?

- Plain model?
- How to train a strong plain model in super-resolution?
  - Direct Training 25.30dB
  - Knowledge Distillation 25.67dB (+0.37dB)
  - RepVGG 24.52dB (-0.78dB)
  - RepVGG-bn-free 25.35dB (+0.05dB)

Solution?

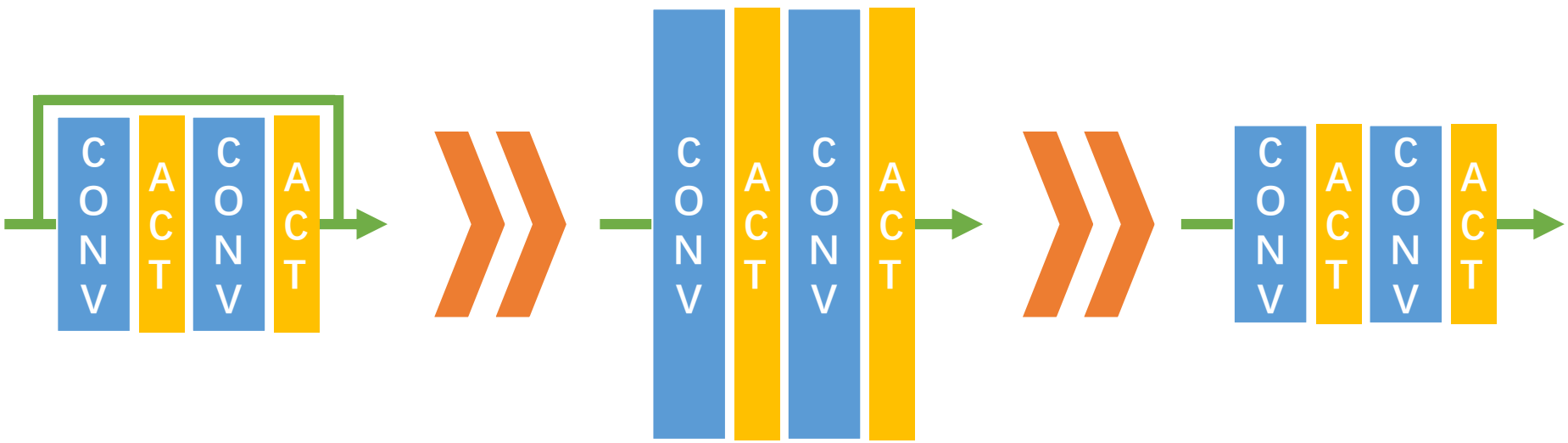
- Plain model?
- How to train a strong plain model in super-resolution?
  - Direct Training — 25.30dB
  - Knowledge Distillation — 25.67dB (+0.37dB)
  - RepVGG — 24.52dB (-0.78dB)
  - RepVGG-bn-free — 25.35dB (+0.05dB)
  - Our Initialization + Direct Training — 25.97dB (+0.67dB)
  - + Knowledge Distillation — 25.99dB (+0.69dB)

Solution



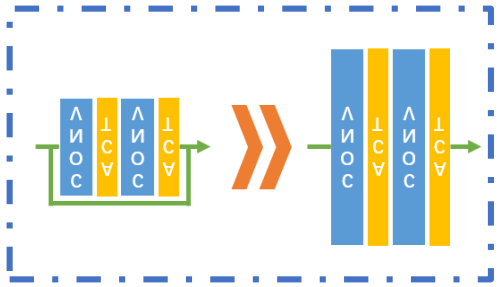
Transform a trained multi-branch teacher model into an equivalent large-size plain model.

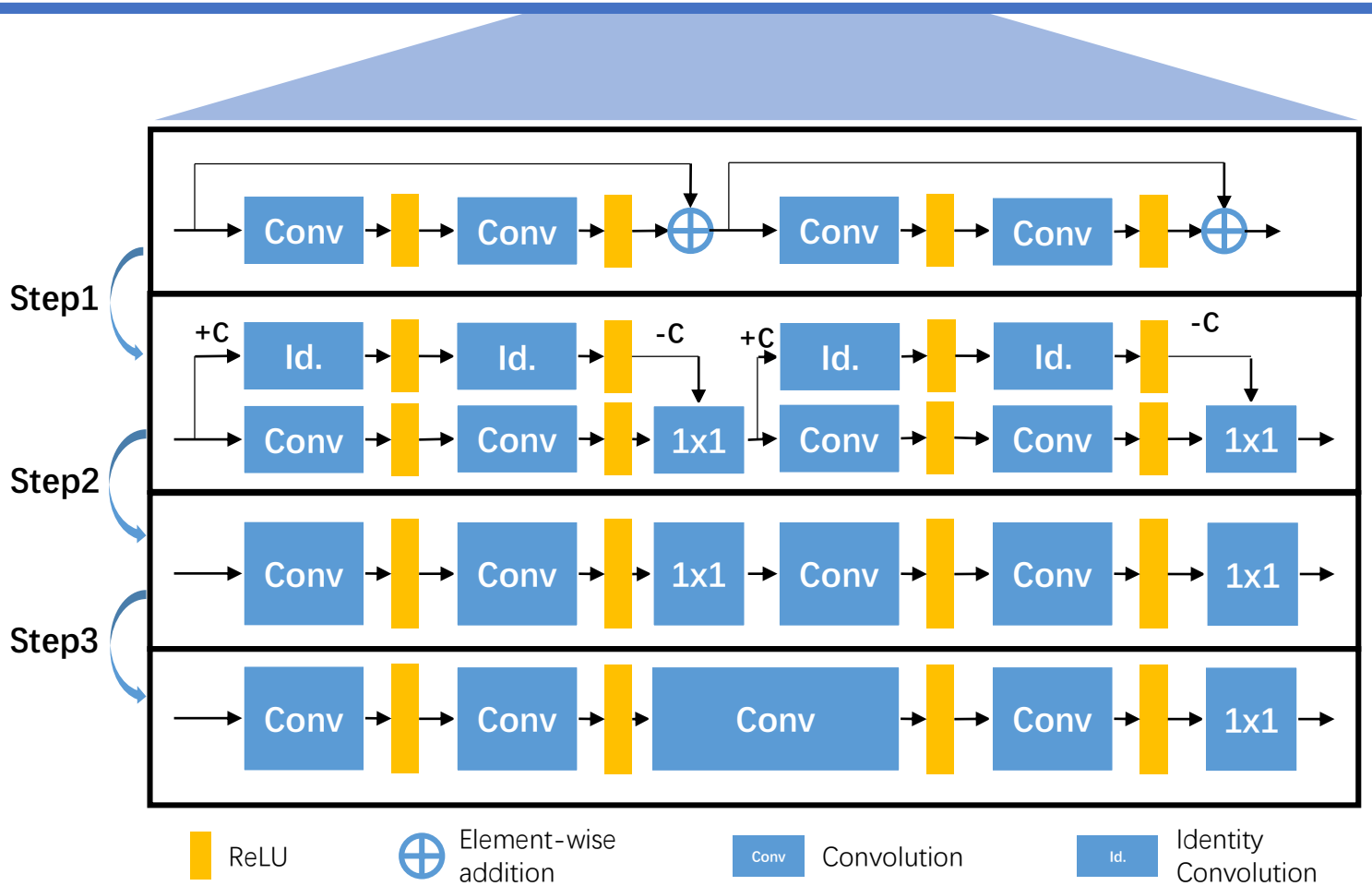
Solution



Use large-size plain model to compute an initialization of the small-size plain model.

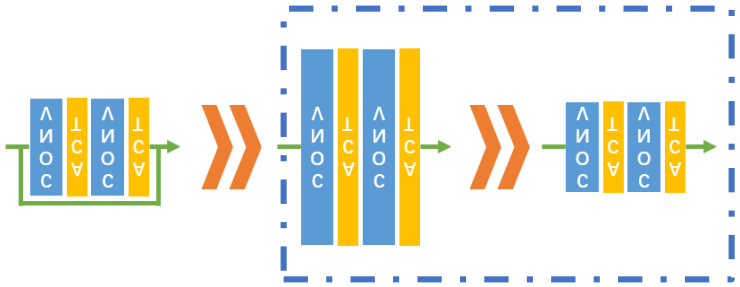
Stage1

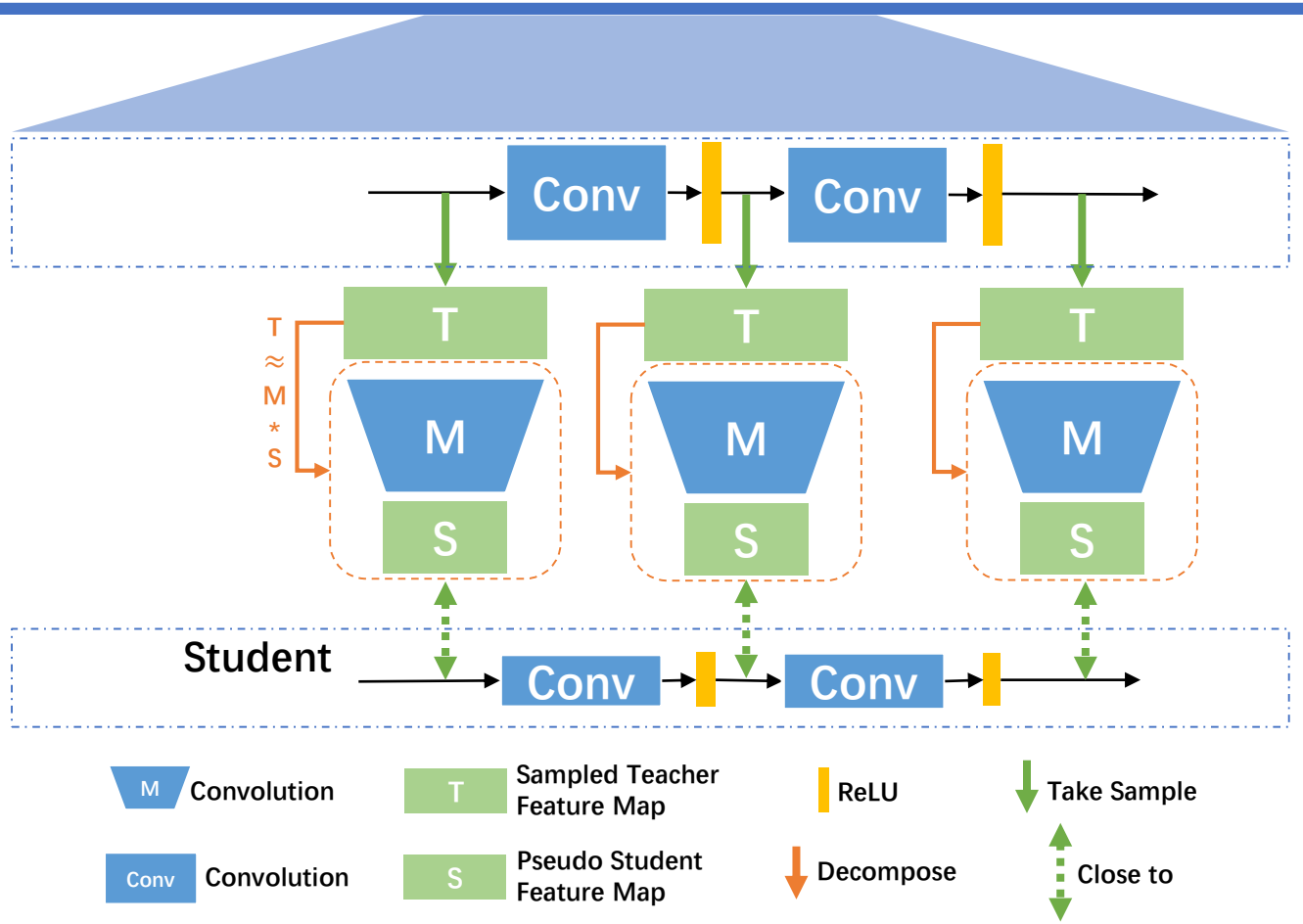




Transforms a trained multi-branch teacher model into an equivalent large-size plain model.

Stage2





Use large-size plain model to compute an initialization of the small-size plain model.

Results

- Performance on standard benchmarks.

Scale	Method	Param.	FLOPs	Runtime	Memory	Set5	Set14	B100	Urban100
						PSNR/SSIM	PSNR/SSIM	PSNR/SSIM	PSNR/SSIM
×2	Bicubic	*	*	*	*	33.66 / 0.9299	30.24 / 0.8688	29.56 / 0.8431	26.88 / 0.8403
	FSRCNN	24K	34G	13ms	**368MB**	37.00 / 0.9558	32.63 / 0.9088	31.53 / 0.8920	29.88 / 0.9020
	Plain-M	1,219K	636G	158ms	420MB	37.99 / 0.9606	33.64 / 0.9184	32.19 / 0.8999	32.10 / 0.9284
	Plain-X$_{\theta=0.05}$	2,224K	1,180G	253ms	718MB	**38.10 / 0.9609**	**33.71 / 0.9187**	**32.24 / 0.9004**	**32.31 / 0.9302**
	SRCNN	68K	144G	60ms	807MB	36.66 / 0.9542	32.45 / 0.9067	31.36 / 0.8879	29.50 / 0.8946
	IDN	577K	393G	138ms	916MB	37.83 / 0.9600	33.30 / 0.9148	32.08 / 0.8985	31.27 / 0.9196
	EDSR	1,370K	713G	199ms	1,306MB	37.99 / 0.9604	33.57 / 0.9175	32.16 / 0.8994	31.98 / 0.9272
	LatticeNet	765K	384G	178ms	1,710MB	38.06 / 0.9607	33.70 / 0.9187	32.20 / 0.8999	32.25 / 0.9288
	RFDN	534K	279G	166ms	1,731MB	38.05 / 0.9606	33.68 / 0.9184	32.16 / 0.8994	32.12 / 0.9278
	IMDN	694K	359G	150ms	1,813MB	38.00 / 0.9605	33.63 / 0.9177	32.19 / 0.8996	32.17 / 0.9283
	CARN	1,592K	503G	199ms	3,210MB	37.76 / 0.9590	33.52 / 0.9166	32.09 / 0.8978	31.92 / 0.9256
	LARAR	548K	384G	211ms	4,098MB	38.01 / 0.9605	33.62 / 0.9183	32.19 / 0.8999	32.10 / 0.9283

Results

- Performance on standard benchmarks.

Scale	Method	Param.	FLOPs	Runtime	Memory	Set5 PSNR/SSIM	Set14 PSNR/SSIM	B100 PSNR/SSIM	Urban100 PSNR/SSIM
×3	Bicubic	*	*	*	*	30.39 / 0.8682	27.55 / 0.7742	27.21 / 0.7385	24.46 / 0.7349
	FSRCNN	24K	30G	6ms	**179MB**	33.18 / 0.9140	29.37 / 0.8240	28.53 / 0.7910	26.43 / 0.8080
	Plain-M	1,219K	285G	68ms	202MB	34.34 / 0.9269	30.31 / 0.8417	29.08 / 0.8048	28.10 / 0.8513
	Plain-X$_{\theta=0.05}$	2,379K	561G	116ms	341MB	**34.47 / 0.9278**	**30.37 / 0.8426**	**29.13 / 0.8062**	**28.29 / 0.8552**
	IDN	577K	239G	61ms	435MB	34.11 / 0.9253	29.99 / 0.8354	28.95 / 0.8013	27.42 / 0.8359
	RFDN	541K	125G	72ms	784MB	34.41 / 0.9273	30.34 / 0.8420	29.09 / 0.8050	28.21 / 0.8525
	LatticeNet	765K	172G	79ms	787MB	34.40 / 0.9272	30.32 / 0.8416	29.10 / 0.8049	28.19 / 0.8513
	SRCNN	68K	144G	60ms	807MB	32.75 / 0.9090	29.30 / 0.8215	28.41 / 0.7863	26.24 / 0.7989
	IMDN	703K	162G	65ms	822MB	34.36 / 0.9270	30.32 / 0.8417	29.09 / 0.8046	28.17 / 0.8519
	EDSR	1,555K	361G	101ms	1,160MB	34.37 / 0.9270	30.28 / 0.8417	29.09 / 0.8052	28.15 / 0.8527
	CARN	1,592K	268G	102ms	1,950MB	34.29 / 0.9255	30.29 / 0.8407	29.06 / 0.8034	28.06 / 0.8493
	LARAR	594K	256G	144ms	4,092MB	34.36 / 0.9267	30.34 / 0.8421	29.11 / 0.8054	28.15 / 0.8523

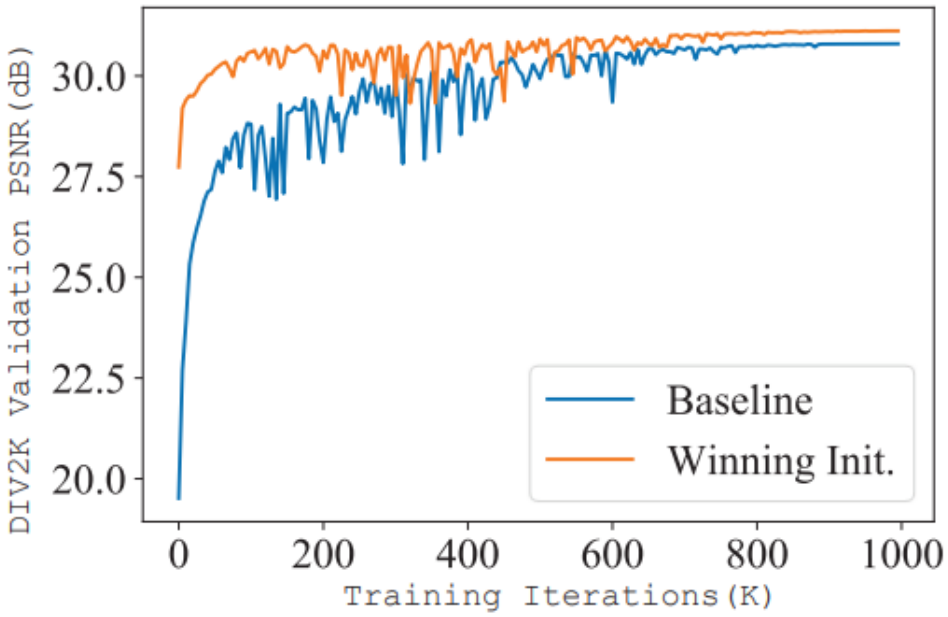
Results

- Performance on standard benchmarks.

Scale	Method	Param.	FLOPs	Runtime	Memory	Set5 PSNR/SSIM	Set14 PSNR/SSIM	B100 PSNR/SSIM	Urban100 PSNR/SSIM
×4	Bicubic	*	*	*	*	28.42 / 0.8104	26.00 / 0.7027	25.96 / 0.6675	23.14 / 0.6577
	FSRCNN	24K	30G	4ms	**111MB**	30.72 / 0.8660	27.61 / 0.7550	26.98 / 0.7150	24.62 / 0.7280
	Plain-M	1,219K	162G	40ms	161MB	32.10 / 0.8938	28.57 / 0.7805	27.55 / 0.7352	25.99 / 0.7830
	Plain-X$_{\theta=0.14}$	1,259K	167G	41ms	168MB	32.14 / 0.8941	28.59 / 0.7810	27.56 / 0.7356	26.05 / 0.7845
	Plain-X$_{\theta=0.05}$	2,541K	337G	70ms	211MB	**32.21 / 0.8950**	**28.63 / 0.7822**	27.60 / **0.7369**	**26.17 / 0.7883**
	IDN	577K	185G	36ms	269MB	31.82 / 0.8903	28.25 / 0.7730	27.41 / 0.7297	25.41 / 0.7632
	RFDN	550K	72G	43ms	457MB	32.24 / 0.8952	28.61 / 0.7819	27.57 / 0.7360	26.11 / 0.7858
	LatticeNet	777K	98G	47ms	473MB	32.18 / 0.8943	28.61 / 0.7812	27.57 / 0.7355	26.14 / 0.7844
	IMDN	715K	92G	38ms	476MB	32.21 / 0.8948	28.58 / 0.7811	27.56 / 0.7353	26.04 / 0.7838
	SRCNN	68K	143.6G	60ms	807MB	30.48 / 0.8628	27.50 / 0.7513	26.90 / 0.7101	24.52 / 0.7221
	EDSR	1,518K	257G	74ms	1,108MB	32.09 / 0.8938	28.58 / 0.7813	27.57 / 0.7357	26.04 / 0.7849
	CARN	1,592K	205G	76ms	1,554MB	32.13 / 0.8937	28.60 / 0.7806	27.58 / 0.7349	26.07 / 0.7837
	LARAR	659K	211G	122ms	4,090MB	32.15 / 0.8944	28.61 / 0.7818	**27.61** / 0.7366	26.14 / 0.7871

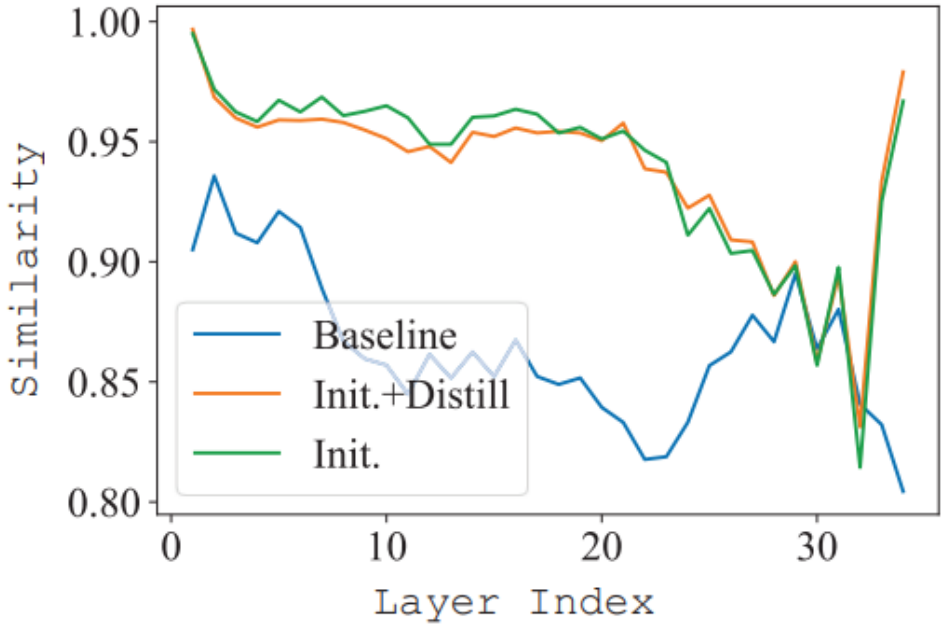
Results

- Training Curve



converges faster and achieves higher accuracy

- Similarity



knowledge is successfully transferred to the student

Thank you!