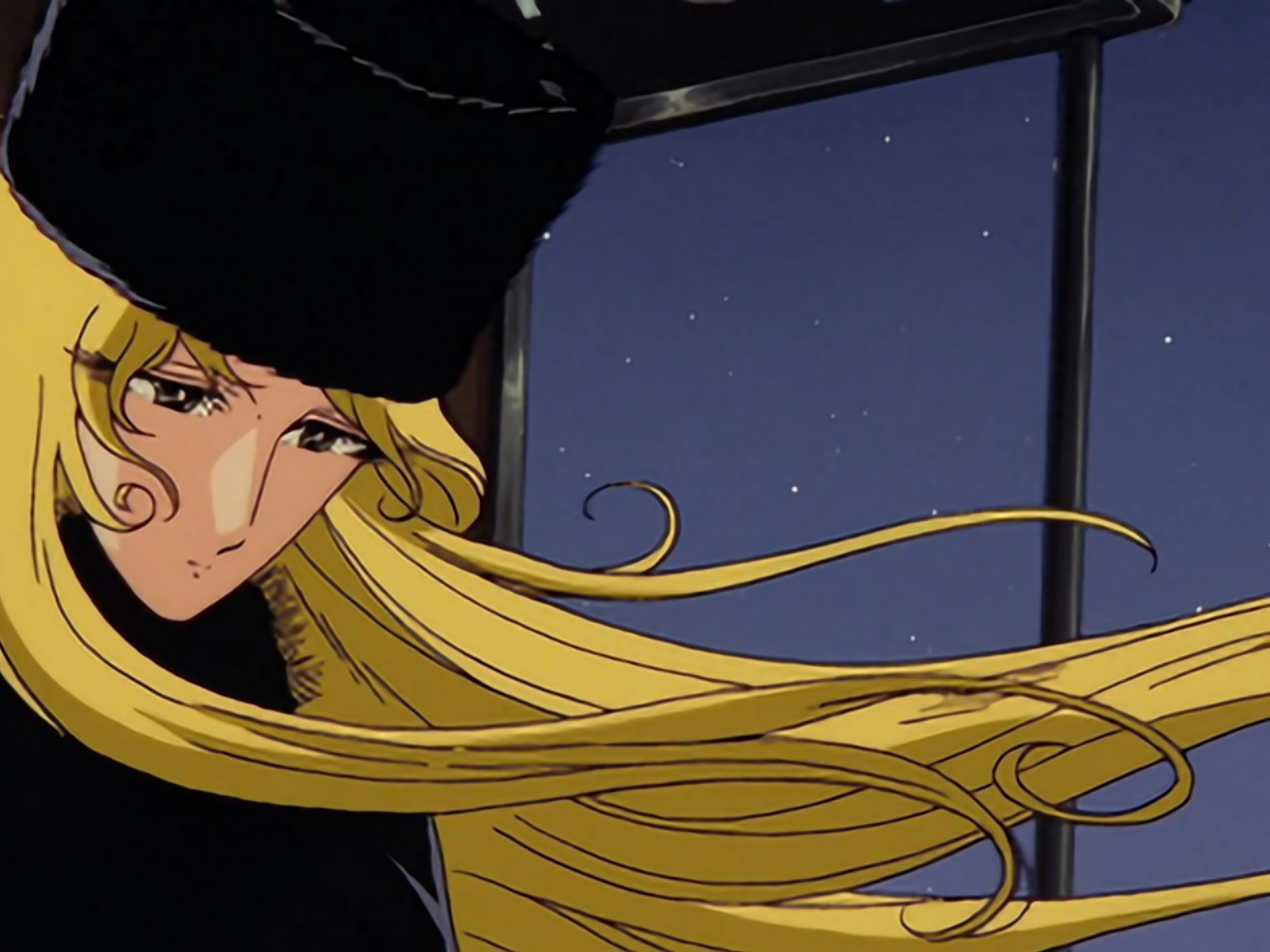
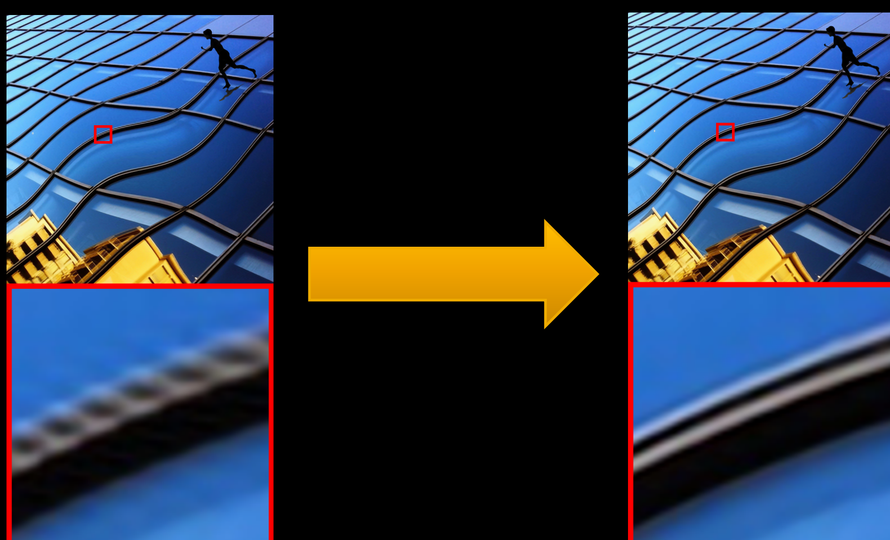


AI KOREA

Deep Image Super-Resolution

Jiwon Kim, Jung Kwon Lee, Kyoung Mu Lee
Seoul National University

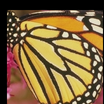


ComputerVisionLab
Seoul National University

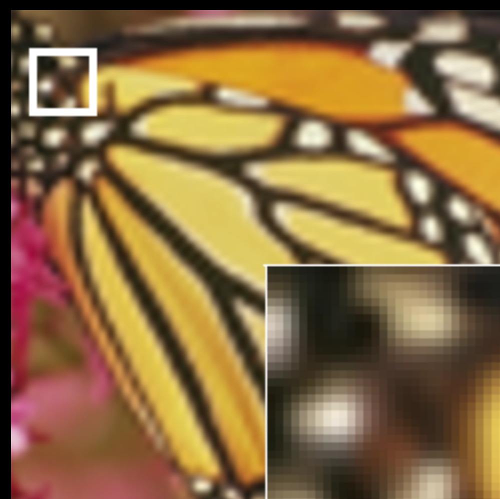
What is Super-Resolution?

We want images BIG & SHARP!

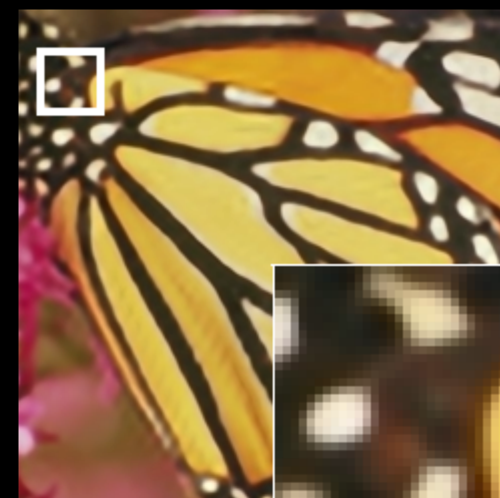
Sharp but Tiny



Big but Blurry *



Big & Sharp

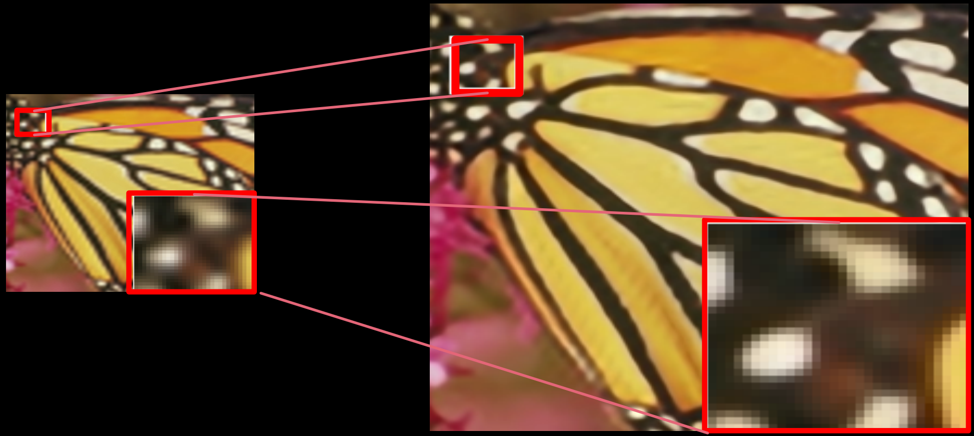


Super-Resolution

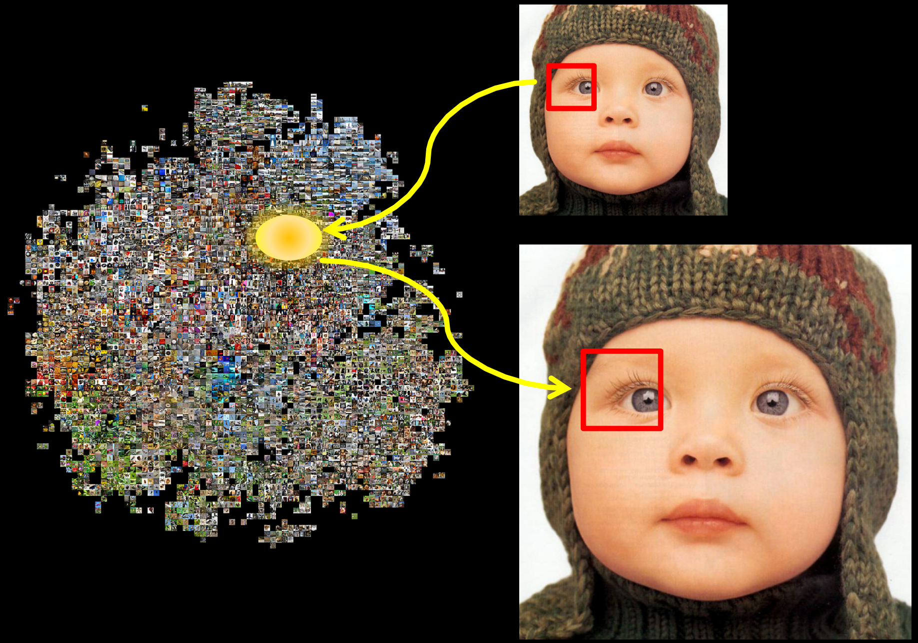
* Most image viewers

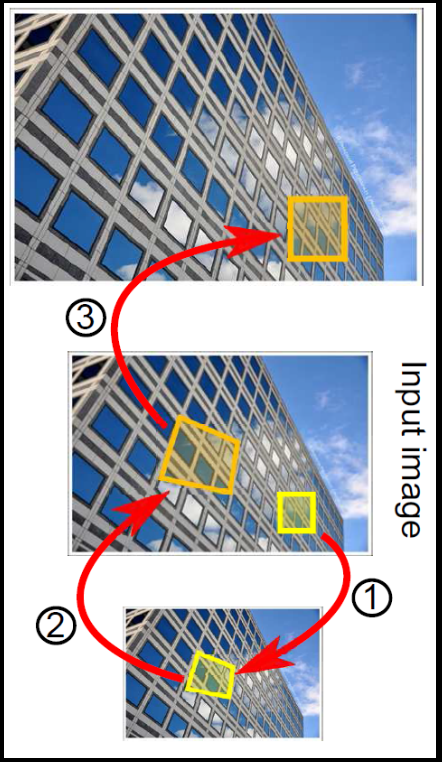
Single-Image Super-Resolution

- Very difficult
  - High-frequency info.(i.e. edges, textures) are lost.
  - Severely ill-posed inversion problem (many possible solutions)
- Algorithms exploit contextual information
  - Often use a low-resolution patch (window of pixels) to predict a high-resolution pixel (window center)



Previous SISR Methods





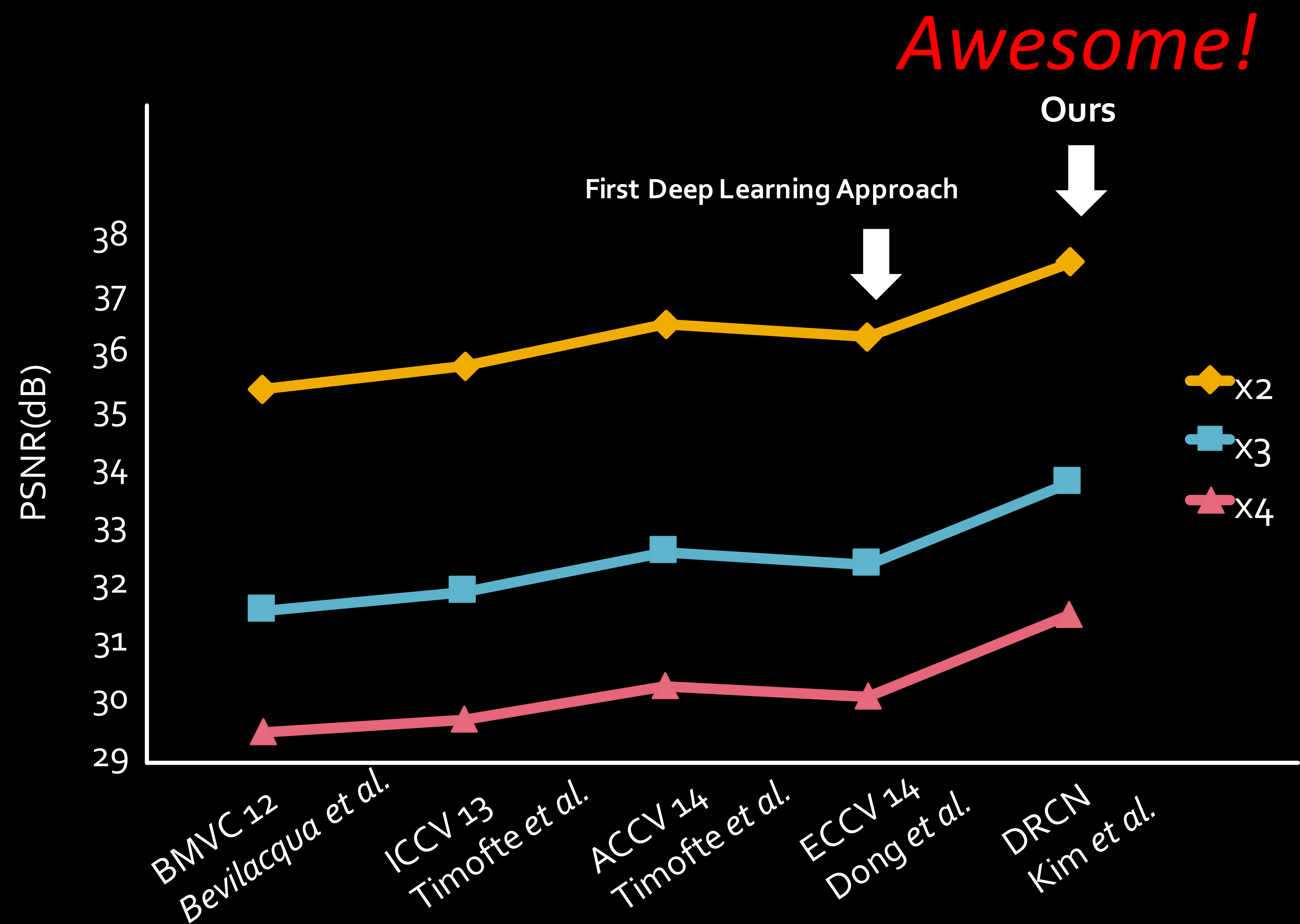
Neighbor Embedding & Dictionary Learning

Self-Similarity

Many algorithms are slow & performances are unsatisfactory

Deep Learning to the Rescue

Dramatic Improvement



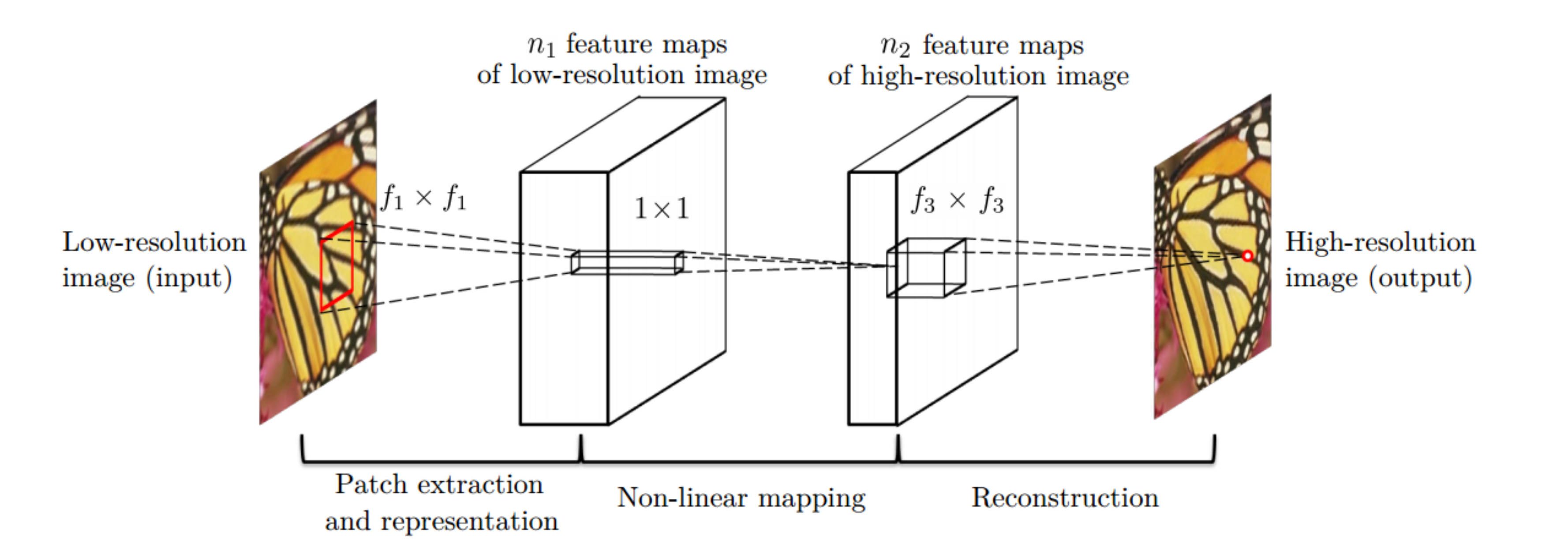
Supervised Learning (Regression)

- Dataset
  - 91 clean natural images (+ 200 BSDS dataset)

- Input – Output
  - Input : interpolated low resolution image (patch)
  - Output : high resolution image (patch)
  - Typically uses bi-cubic interpolation

First Deep Learning Method - SRCNN

- Very shallow, 3-layer CNN
  - Very simple and easy to implement
  - Can be trained in an end-to-end manner
  - Good performance
- Popular (cited 107 times as of today)
  - Waifu2x implementation has 5091 stars on GitHub



Shortcomings of SRCNN

- 1. SRCNN relies on the context of small image regions (13x13)
- 2. The SRCNN network only works for a single scale
- 3. Training of SRCNN converges too slowly (learning rate is 10^{-5}).

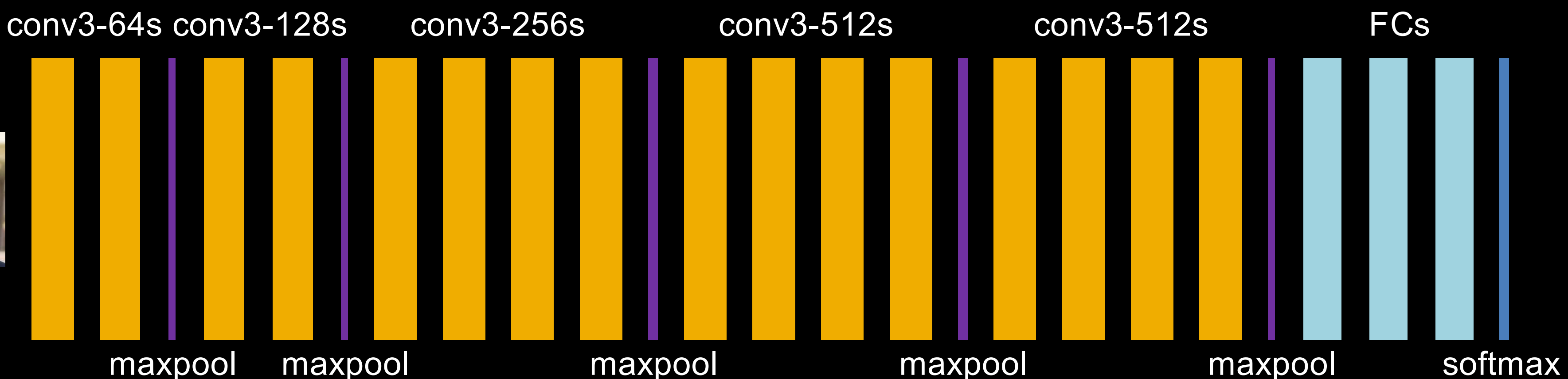
Our New Method
1. Very Deep Super Resolution

Related Work - VGG-Net

- Very Deep Convolutional Neural Network
  - Image recognition task
  - Influenced many CNN-based approaches

Q : Who is this ?

VGG-19 Network



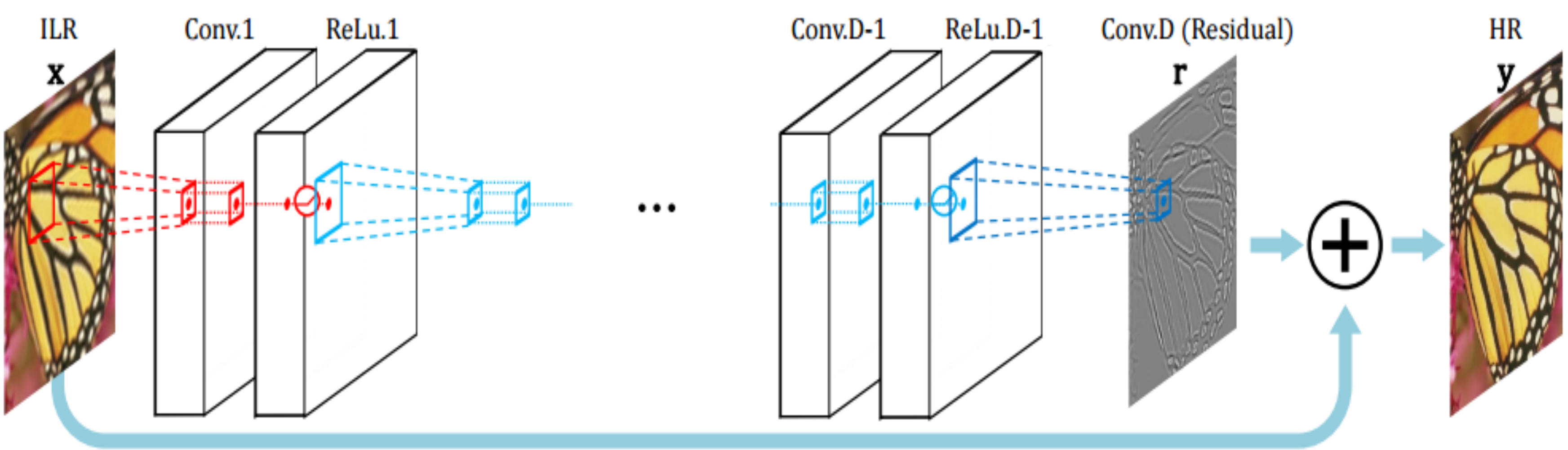
A : Jiwon Kim

VDSR, VGG-like Very Deep Network

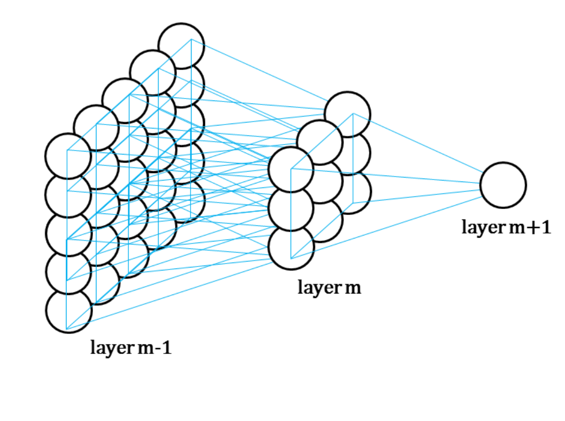
- *Like VGG-Net, exploits very deep CNN*
  - *Up to 20 layers*
  - *With deeper layers, network can see larger part of image*
- *Aggregates all the upscale factors into one single model*
  - *Different scale factors help each other*
- *Can be trained 10^4 faster, with residual*

Model

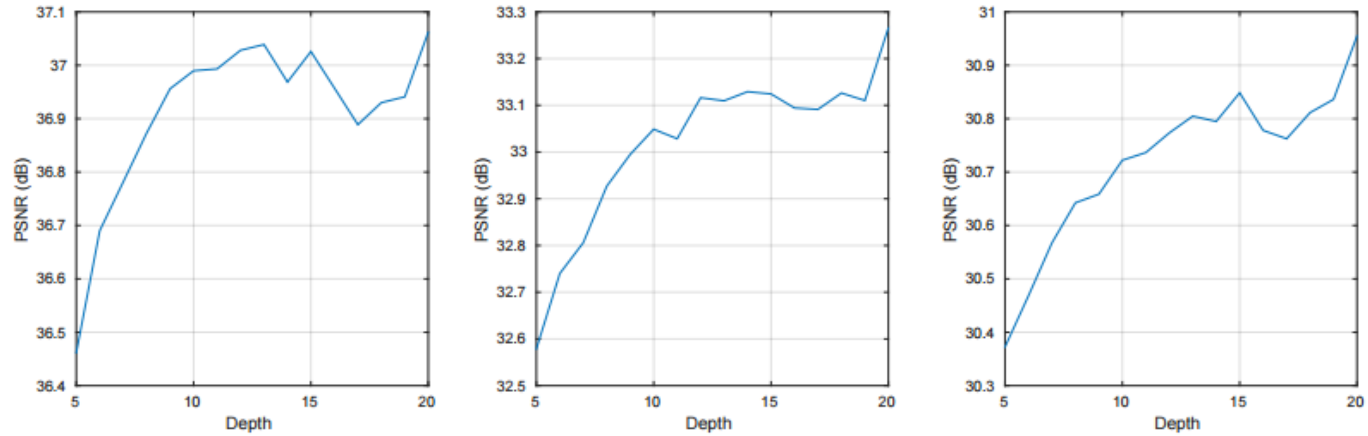
- Uses only 3x3 filters in convolutional layers
- Learn residual only, high frequency part of image
- No dimension reduction such as pooling



The Deeper, the Better

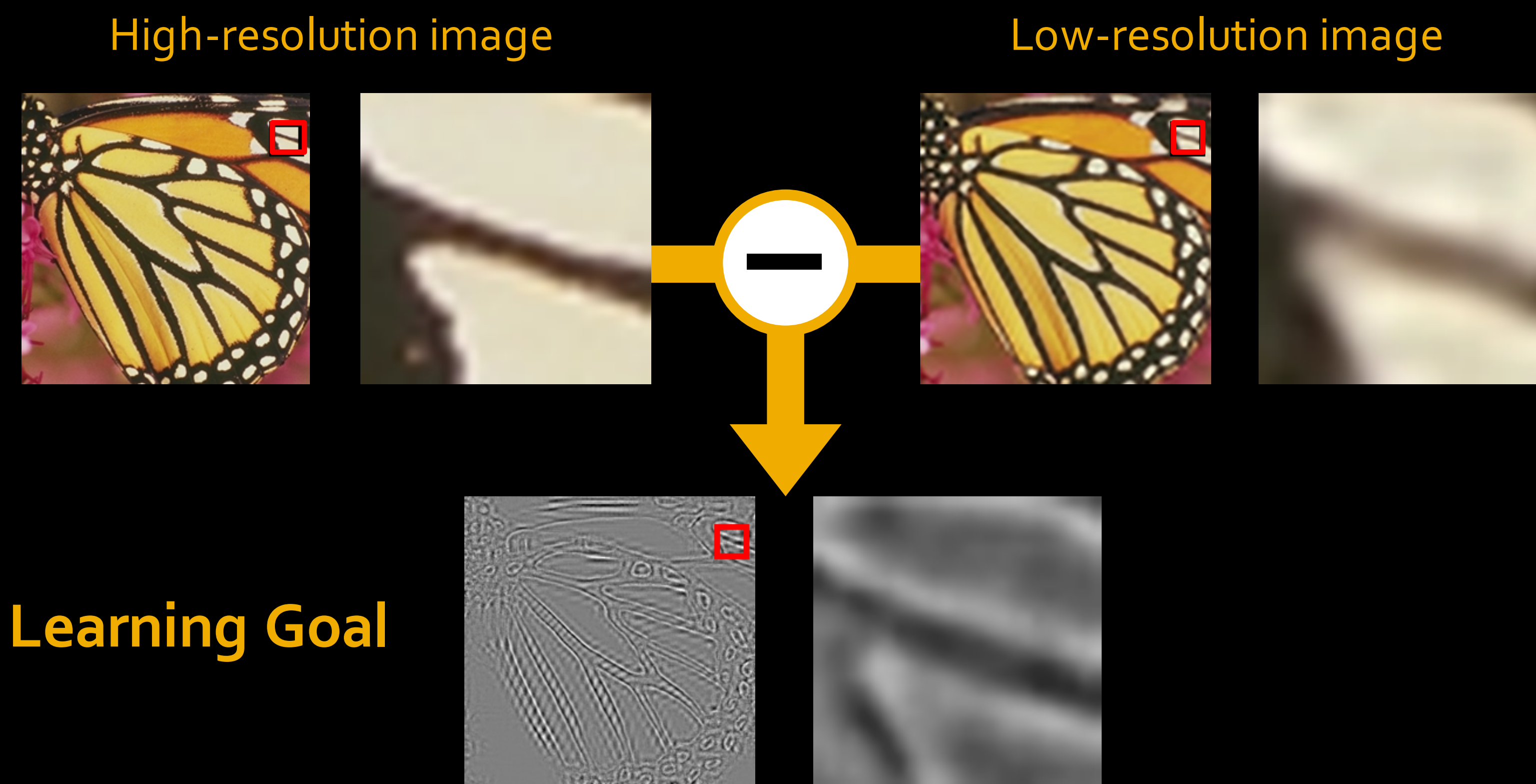


For depth D network,
the receptive field has size
$(2D + 1) \times (2D + 1)$



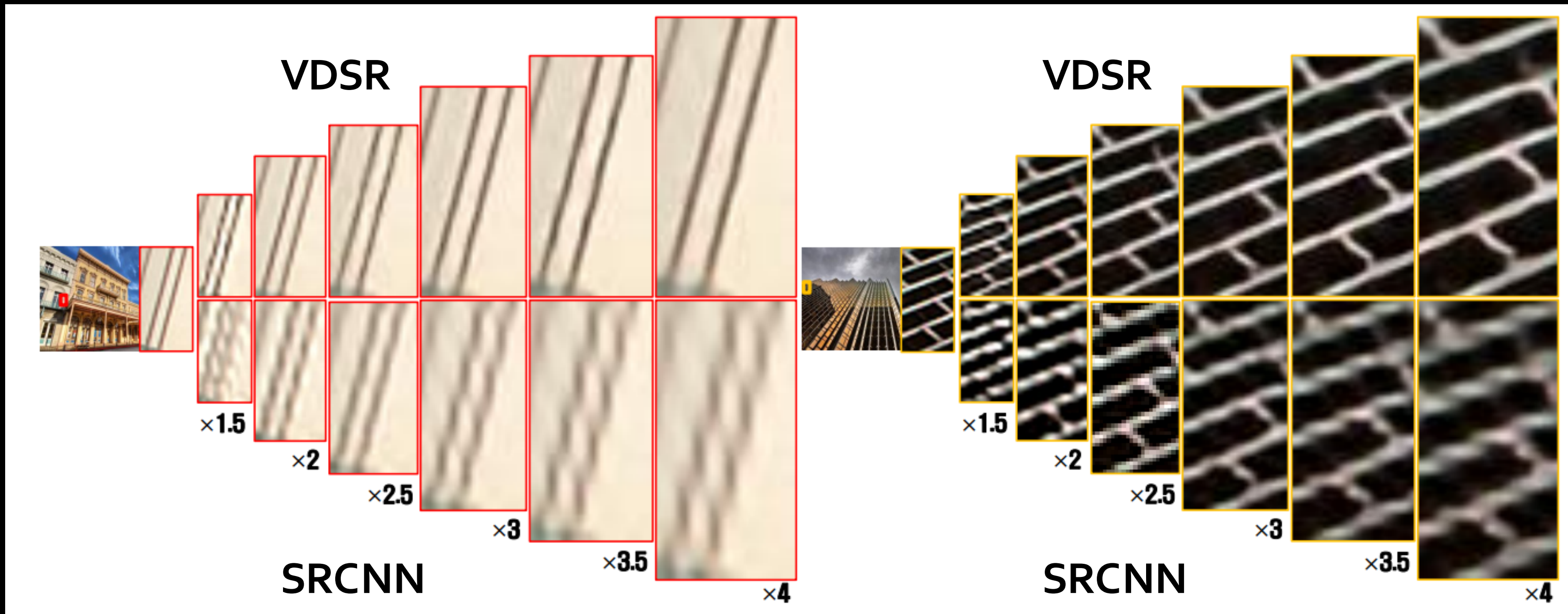
(a) Test Scale Factor 2 (b) Test Scale Factor 3 (c) Test Scale Factor 4

Residual Learning



Only Learns the difference → reduces the problem capacity
Enables much FASTER learning

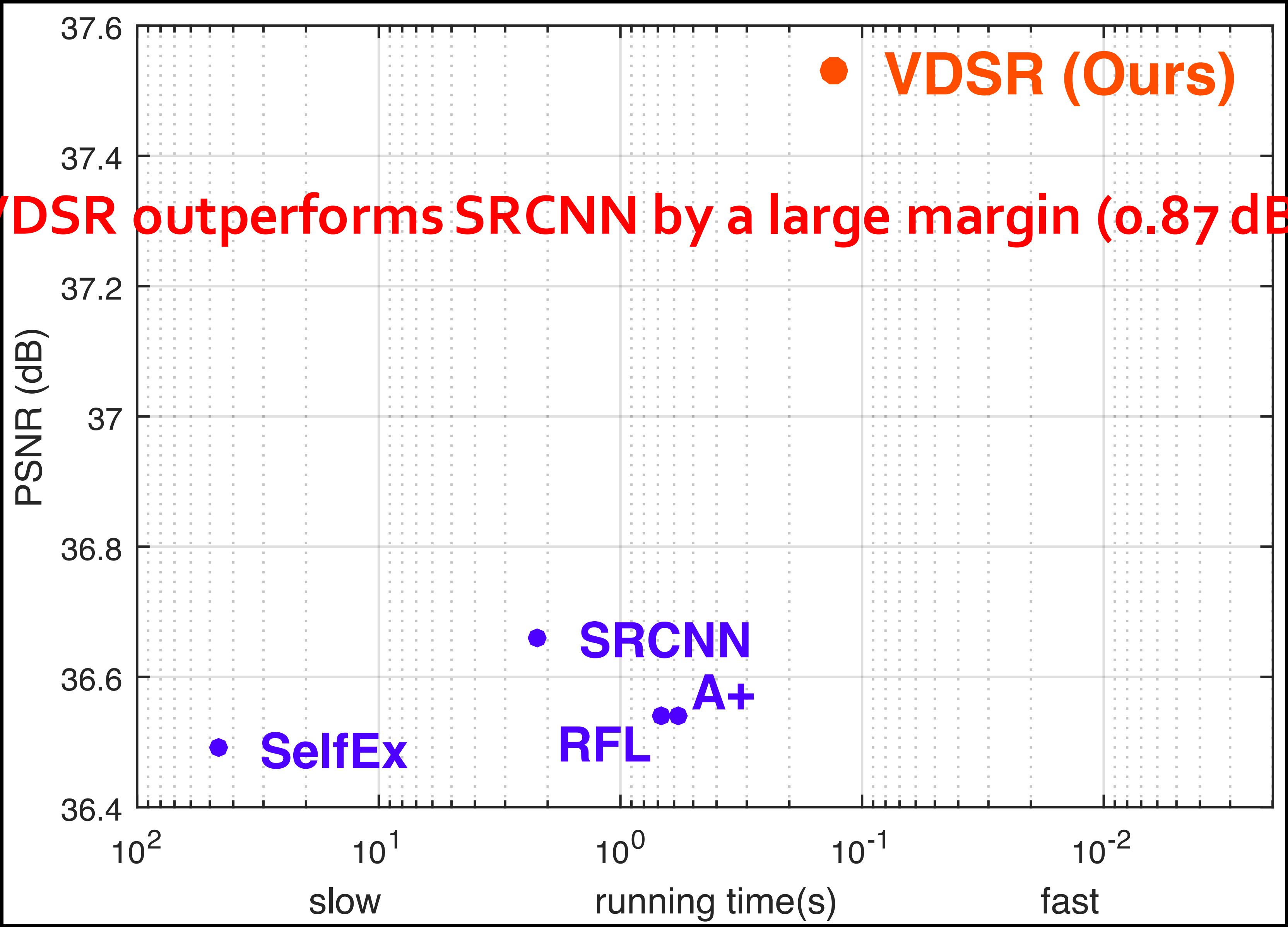
Single Model for Multiple scales



Test / Train	×2	×3	×4	×2,3	×2,4	×3,4	×2,3,4	Bicubic
×2	37.10	30.05	28.13	37.09	37.03	32.43	37.06	33.66
×3	30.42	32.89	30.50	33.22	31.20	33.24	33.27	30.39
×4	28.43	28.73	30.84	28.70	30.86	30.94	30.95	28.42

Interestingly, we observe that training multiple scales boosts the performance for large scales.

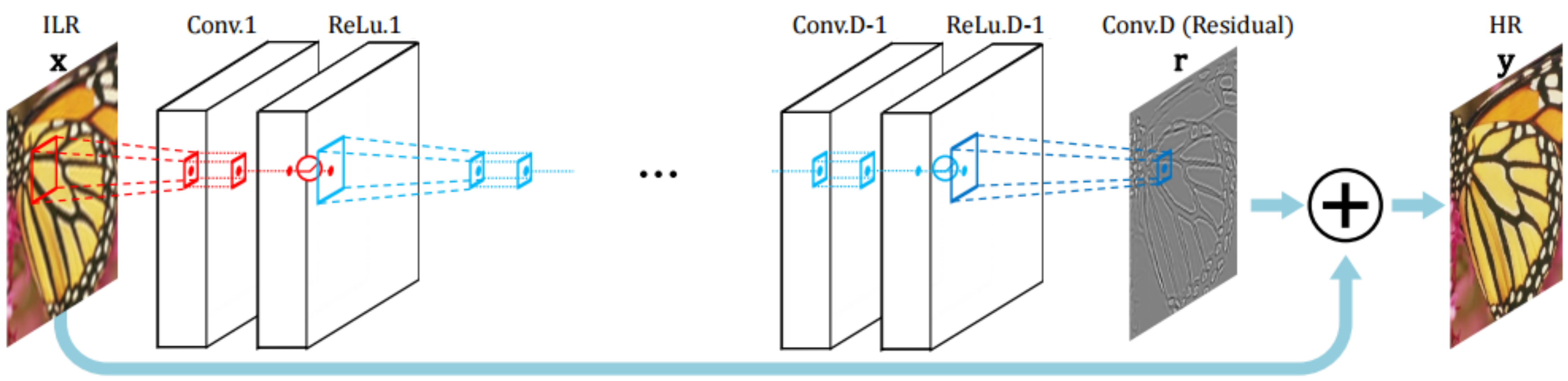
Experimental Result



One Important Observation

- In VDSR, we observed that

Convolution layers exactly have the SAME structures reminding us of RECURSIONs



Our New Method

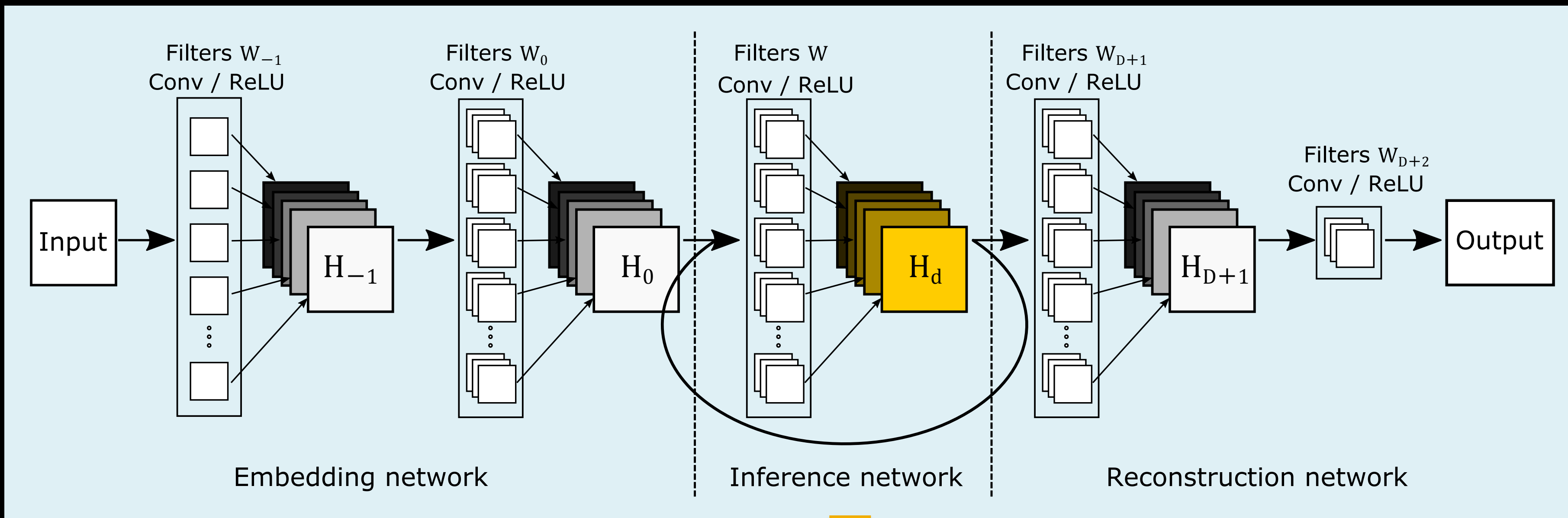
2. Deeply-Recursive Convolutional Network

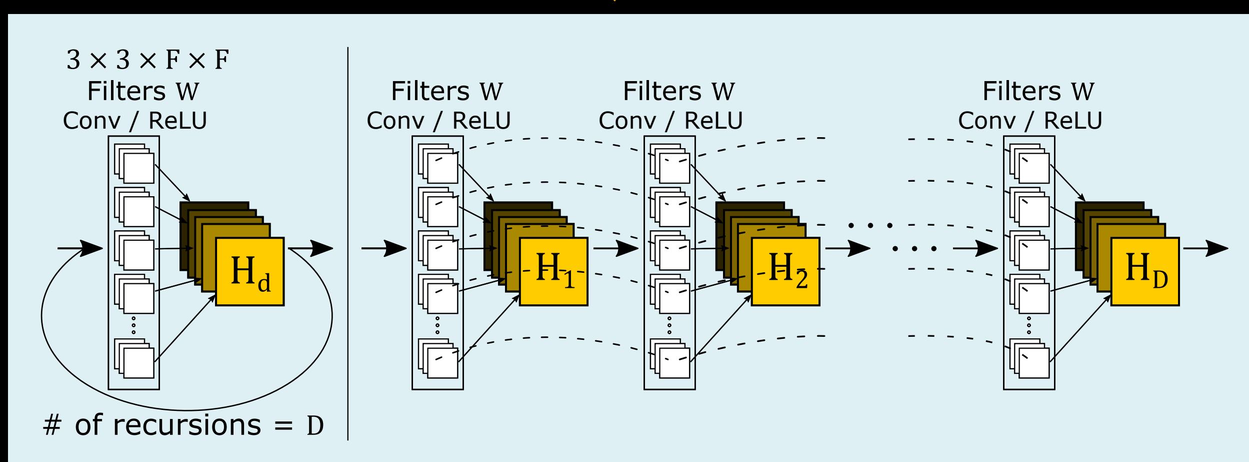
DRCN,
Deeply-Recursive Convolutional Network

- *New method using a deeply-recursive convolutional network (DRCN).*
  - *Very deep recursive layer (up to 16 recursions) – very large receptive field (41x41 vs 13x13 of SRCNN)*
  - *Increasing recursion depth can improve performance without introducing new parameters for additional convolutions.*

- *Learning a DRCN is very hard due to exploding/vanishing gradients. To ease the difficulty of training, we propose two extensions:*
  - *Recursive-supervision*
  - *Skip-connection*

Basic Model: Overview



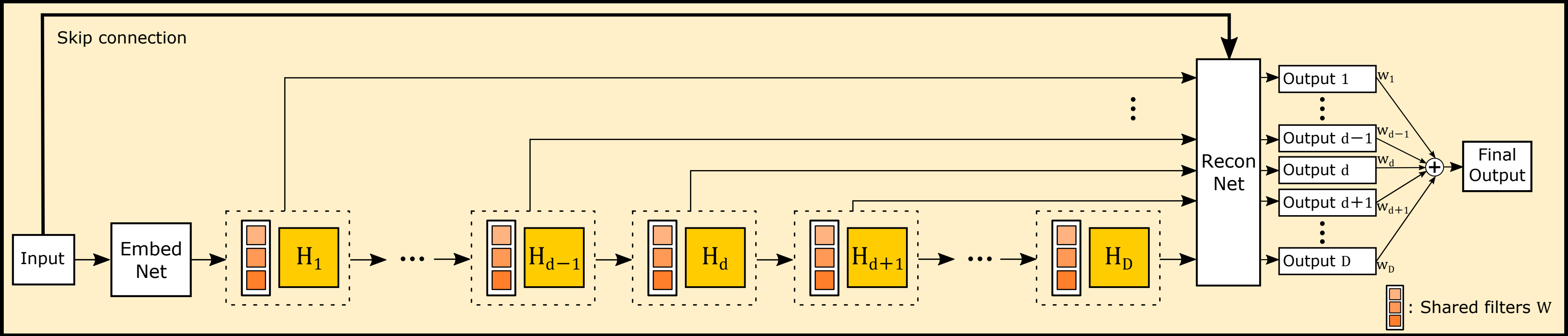


Basic Model: Weakness

- *Exploding/vanishing gradients make the training difficult. Basic model does not converge.*

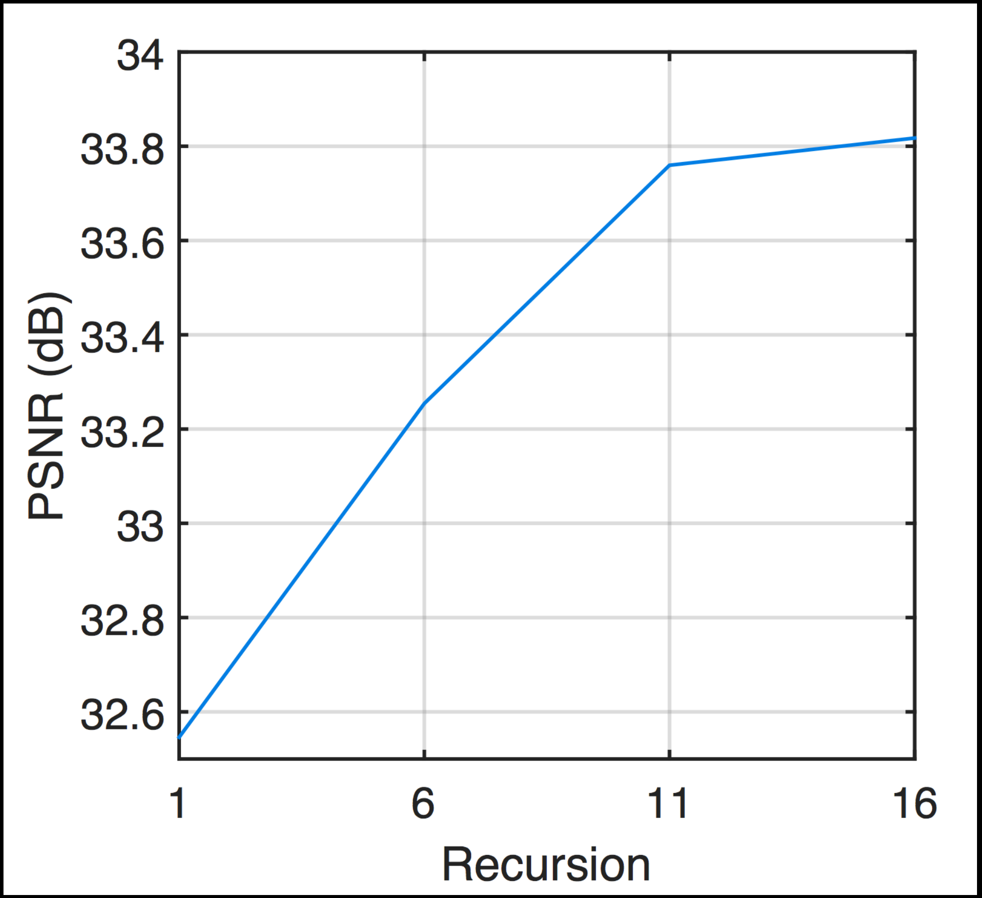
- *To ease the difficulty, we propose two extensions*
  - *Recursive-supervision: all recursions are supervised*
  - *Skip-connection: input directly goes into reconstruction net*

Advanced Model



- *Early recursions are also supervised.*
- *As outputs reconstructed from all depths are ensembled, cherry-picking the optimal depth is not required.*
- *Exact copy of input image is not lost during recursions.*

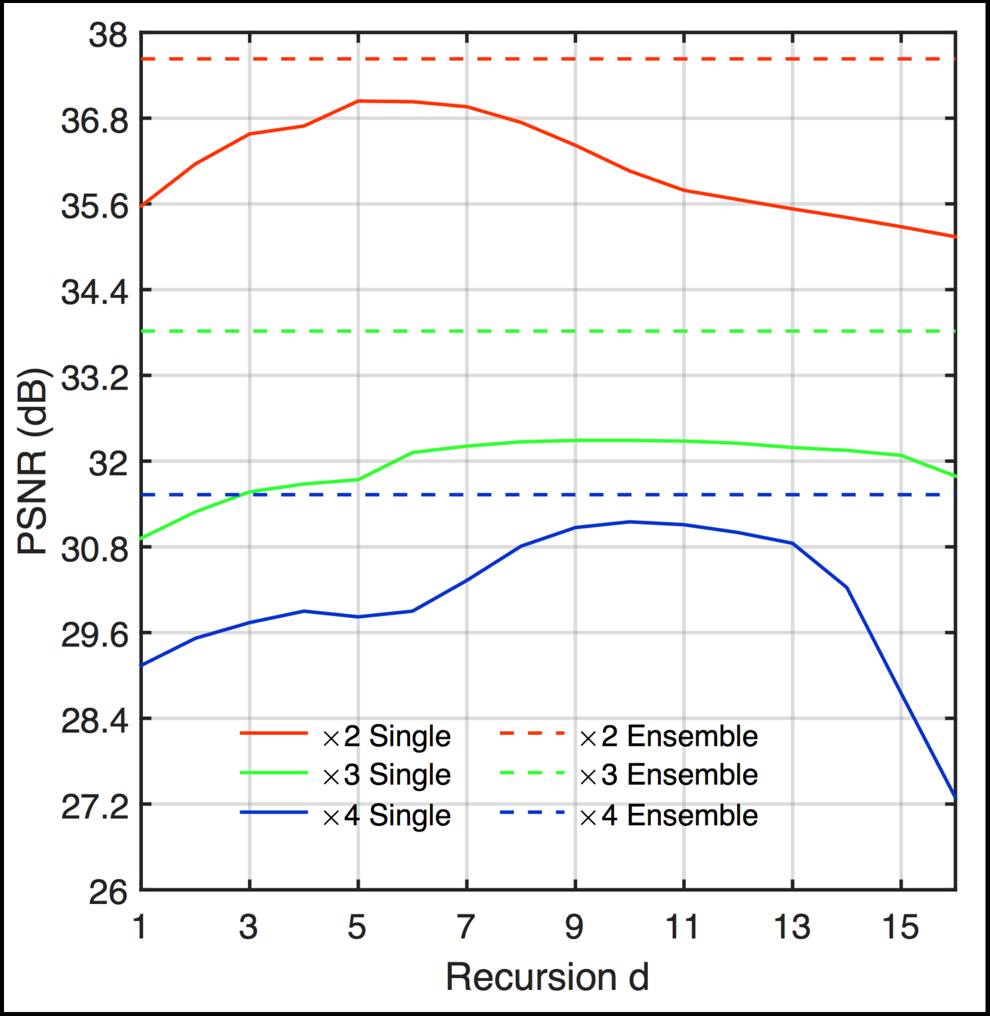
Study of Advanced Model – Recursion effect



Recursion versus performance for the scale factor ×3 on the dataset *Set5*.

More recursions yielding larger receptive fields lead to better performances.

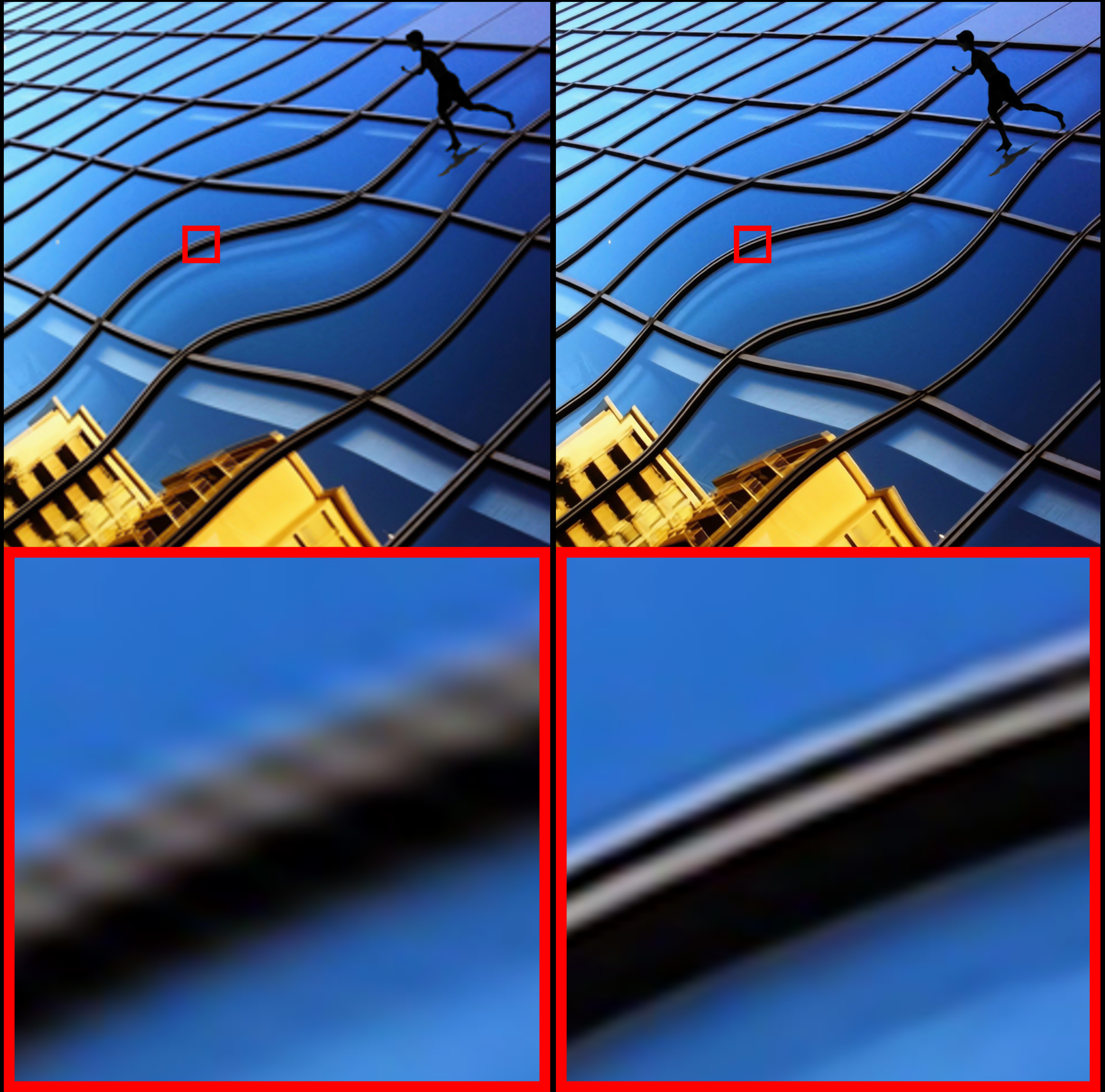
Study of Advanced Model – Ensemble Effect



Prediction made from intermediate recursions are evaluated. There is no single recursion depth that works the best across all scale factors. Ensemble of intermediate predictions significantly improves performance.

Results

x4



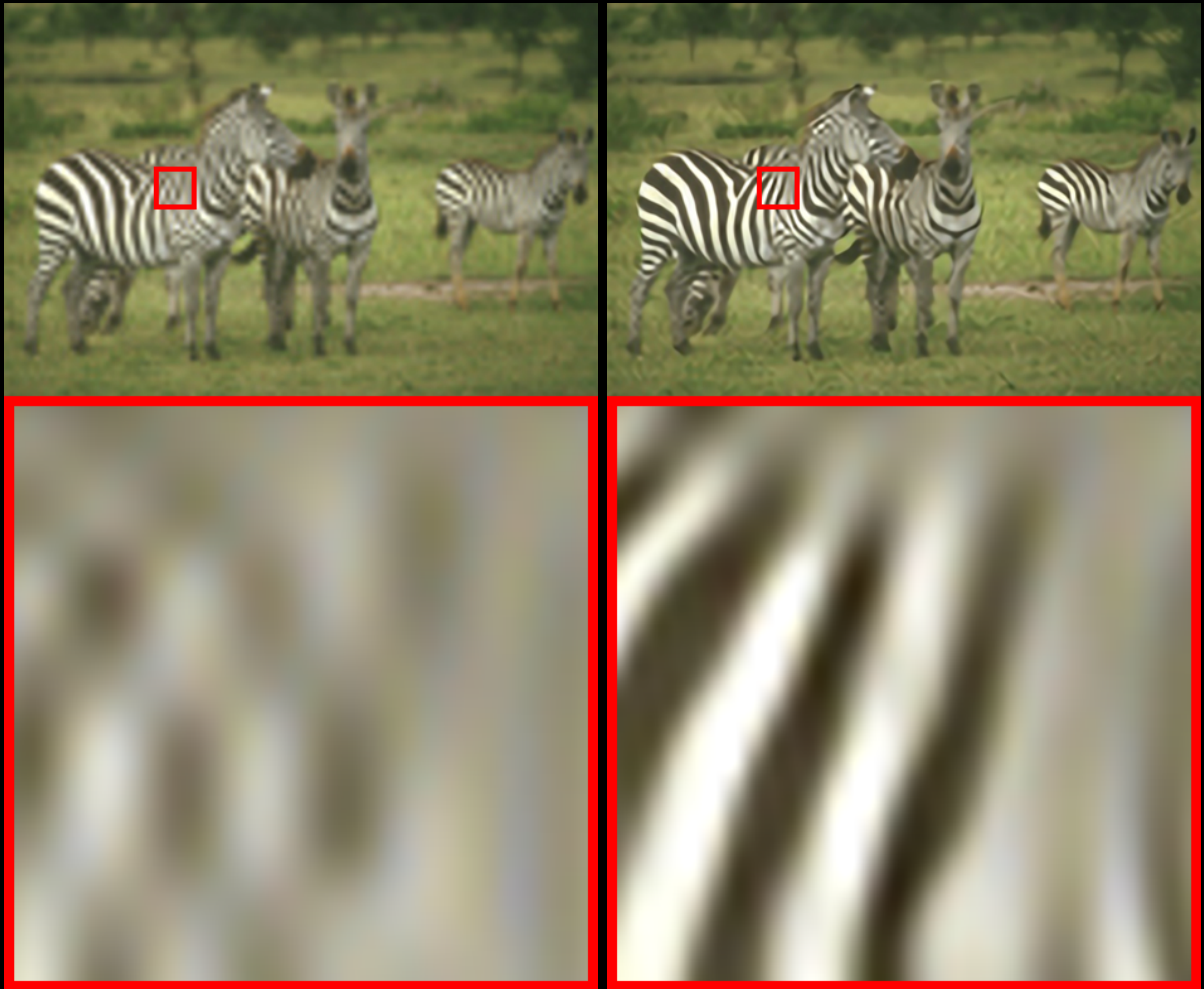
x4

x4

x4

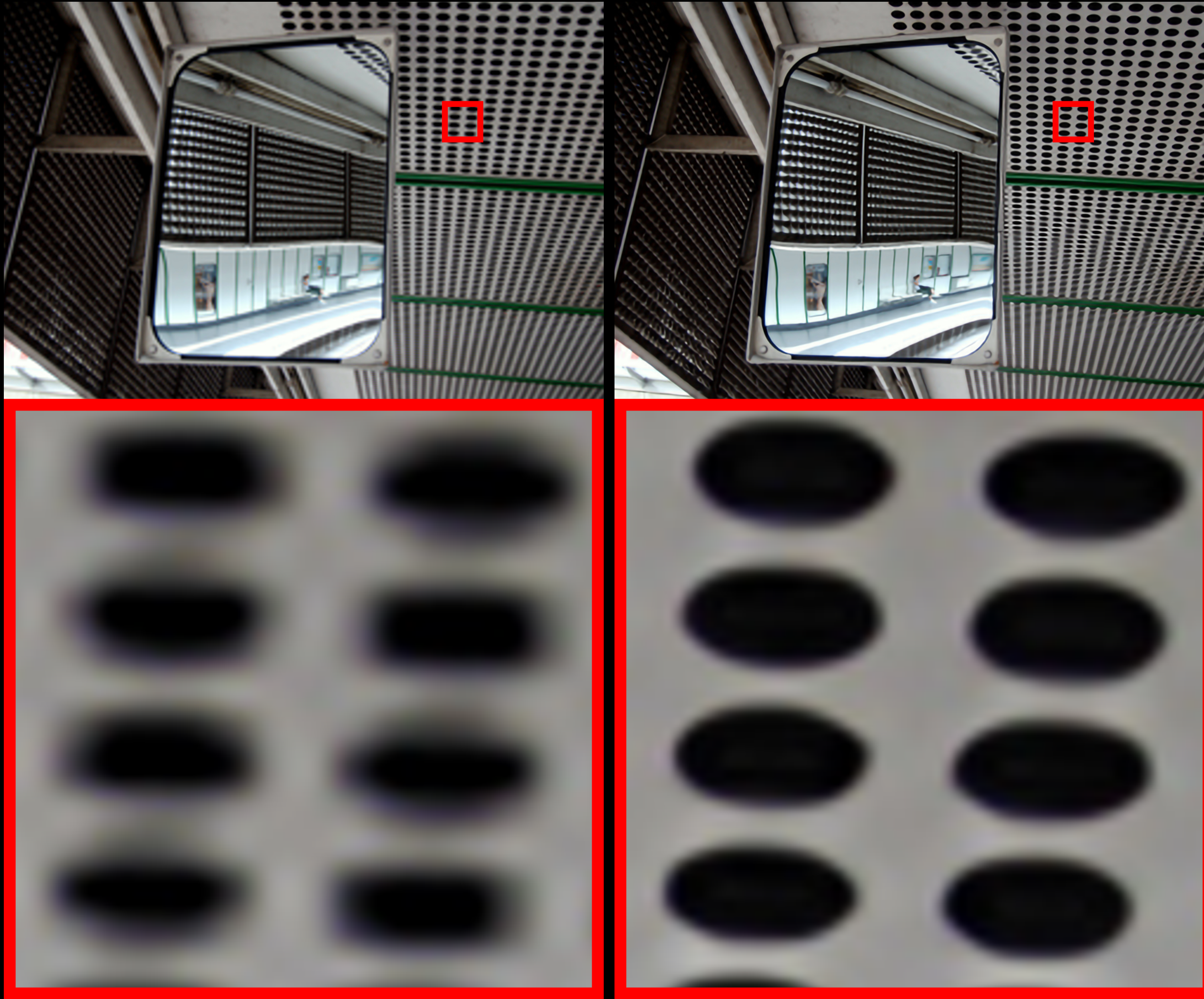

x4

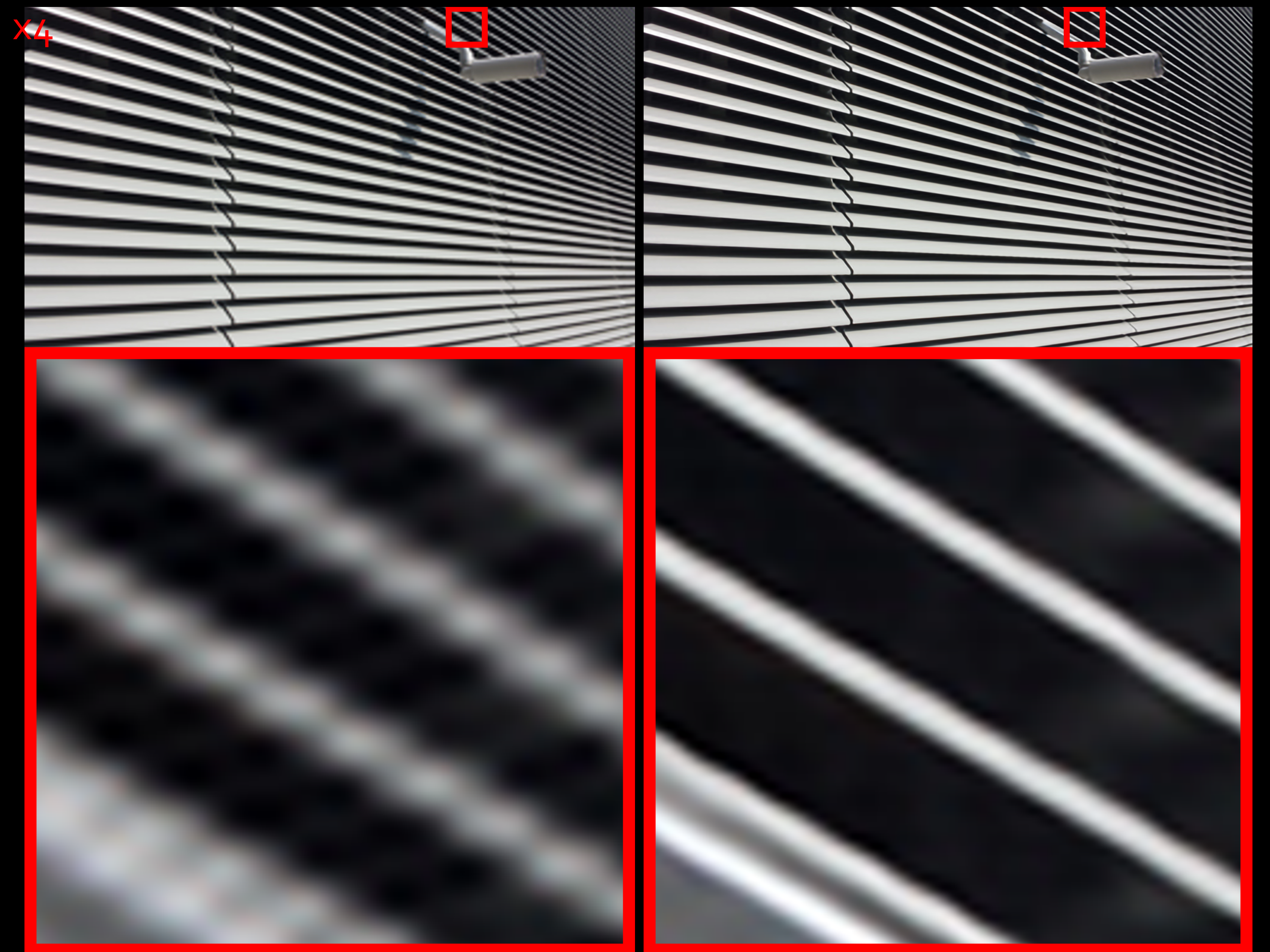
x4



x4

x4




x4

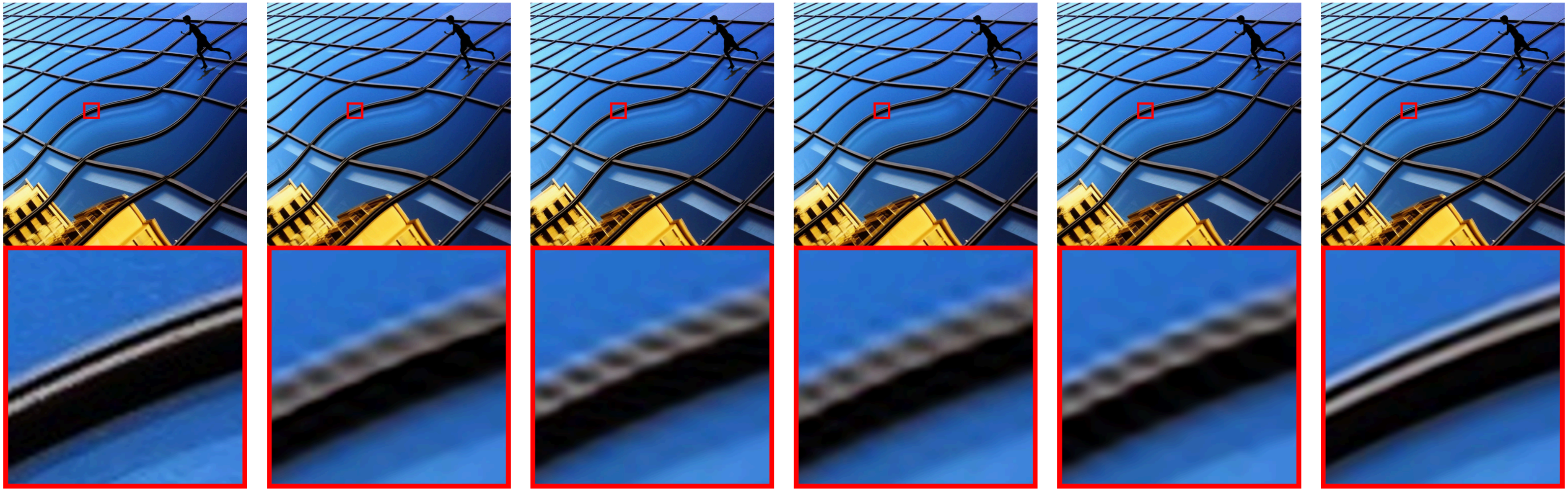
x4


How to Do Everything with
PowerPoint 2002
Create effective multimedia presentations
Make your point in a meaningful, memorable way
Deliver your show on-screen or on the Web
Ellen Finkelstein
OSBORNE

x3


How to Do Everything with
PowerPoint 2002
Create effective multimedia presentations
Make your point in a meaningful, memorable way
Deliver your show on-screen or on the Web
Ellen Finkelstein
OSBORNE
How to Do Everything with
PowerPoint 2002
Create effective multimedia presentations
Make your point in a meaningful, memorable way
Deliver your show on-screen or on the Web
Ellen Finkelstein
OSBORNE

Experimental Results (scale factor 4)

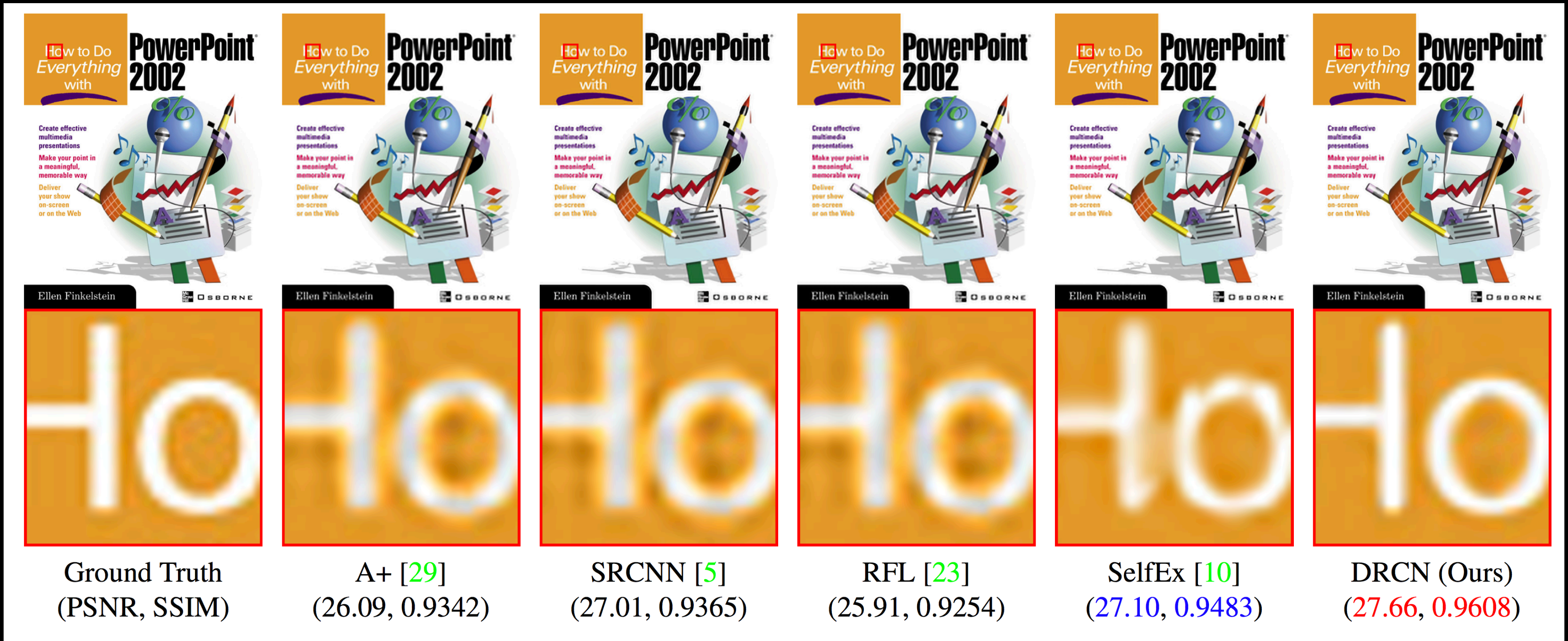


Ground Truth (PSNR, SSIM)	A+ [29] (29.83, 0.9102)	SRCNN [5] (29.97, 0.9092)	RFL [23] (29.61, 0.9026)	SelfEx [10] (30.73, 0.9193)	DRCN (Ours) (32.17, 0.9350)

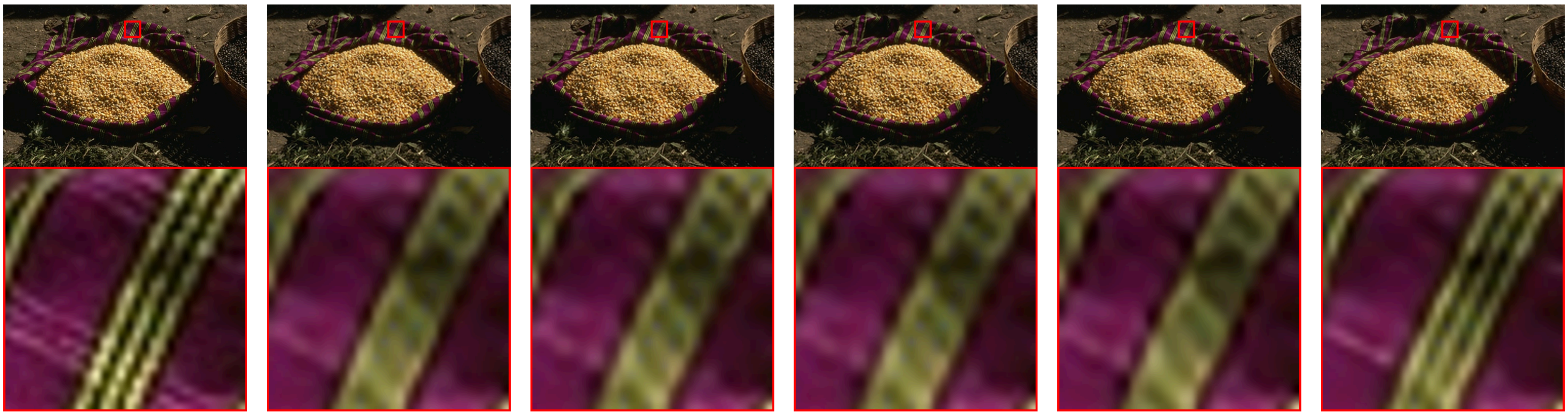
Experimental Results (scale factor 4)

Ground Truth	A+ [29]	SRCNN [5]	RFL [23]	SelfEx [10]	DRCN (Ours)
(PSNR, SSIM)	(23.53, 0.6977)	(23.79, 0.7087)	(23.53, 0.6943)	(23.52, 0.7006)	(24.36, 0.7399)

Experimental Results (scale factor 3)



Experimental Results (scale factor 4)



Ground Truth	A+ [29]	SRCNN [5]	RFL [23]	SelfEx [10]	DRCN (Ours)
(PSNR, SSIM)	(24.24, 0.8176)	(24.48, 0.8267)	(24.24, 0.8137)	(24.16, 0.8145)	(24.76, 0.8385)

Benchmark

Dataset	Scale	Bicubic PSNR/SSIM	A+ [29] PSNR/SSIM	SRCNN [5] PSNR/SSIM	RFL [23] PSNR/SSIM	SelfEx [10] PSNR/SSIM	DRCN (Ours) PSNR/SSIM
Set5	×2	33.66/0.9299	36.54/0.9544	36.66/0.9542	36.54/0.9537	36.49/0.9537	37.63/0.9588
	×3	30.39/0.8682	32.58/0.9088	32.75/0.9090	32.43/0.9057	32.58/0.9093	33.82/0.9226
	×4	28.42/0.8104	30.28/0.8603	30.48/0.8628	30.14/0.8548	30.31/0.8619	31.53/0.8854
Set14	×2	30.24/0.8688	32.28/0.9056	32.42/0.9063	32.26/0.9040	32.22/0.9034	33.04/0.9118
	×3	27.55/0.7742	29.13/0.8188	29.28/0.8209	29.05/0.8164	29.16/0.8196	29.76/0.8311
	×4	26.00/0.7027	27.32/0.7491	27.49/0.7503	27.24/0.7451	27.40/0.7518	28.02/0.7670
B100	×2	29.56/0.8431	31.21/0.8863	31.36/0.8879	31.16/0.8840	31.18/0.8855	31.85/0.8942
	×3	27.21/0.7385	28.29/0.7835	28.41/0.7863	28.22/0.7806	28.29/0.7840	28.80/0.7963
	×4	25.96/0.6675	26.82/0.7087	26.90/0.7101	26.75/0.7054	26.84/0.7106	27.23/0.7233
Urban100	×2	26.88/0.8403	29.20/0.8938	29.50/0.8946	29.11/0.8904	29.54/0.8967	30.75/0.9133
	×3	24.46/0.7349	26.03/0.7973	26.24/0.7989	25.86/0.7900	26.44/0.8088	27.15/0.8276
	×4	23.14/0.6577	24.32/0.7183	24.52/0.7221	24.19/0.7096	24.79/0.7374	25.14/0.7510

- Previous methods are very recent (within 1 year) – this area is getting very competitive.
- But ours outperforms by a large margin – thanks to recent advances in very-deep learning.

Conclusion

- Novel method using deeply-recursive convolutions - additional recursion introduces no additional weight parameters (fixed capacity)
- Recursive-supervision and skip-connection are used for better training.
- Can be applied to other image restoration problems (coming soon).
- The state-of-the-art performance

Thank you!