

Progress of Single-Image Super Resolution Using Deep Learning

2019.01.12

이진호

목 차

- 1. SISR?
- 2. SRCNN
- 3. VDSR
- 4. ESPCN
- 5. SRResnet
- 6. EDSR
- 7. RDN
- 8. SRGAN

Single-Image Super Resolution(SISR)?

- Single Image Super-Resolution
 - ILL-posed problem (many candidate solutions)



저해상도 이미지
(Low resolution)

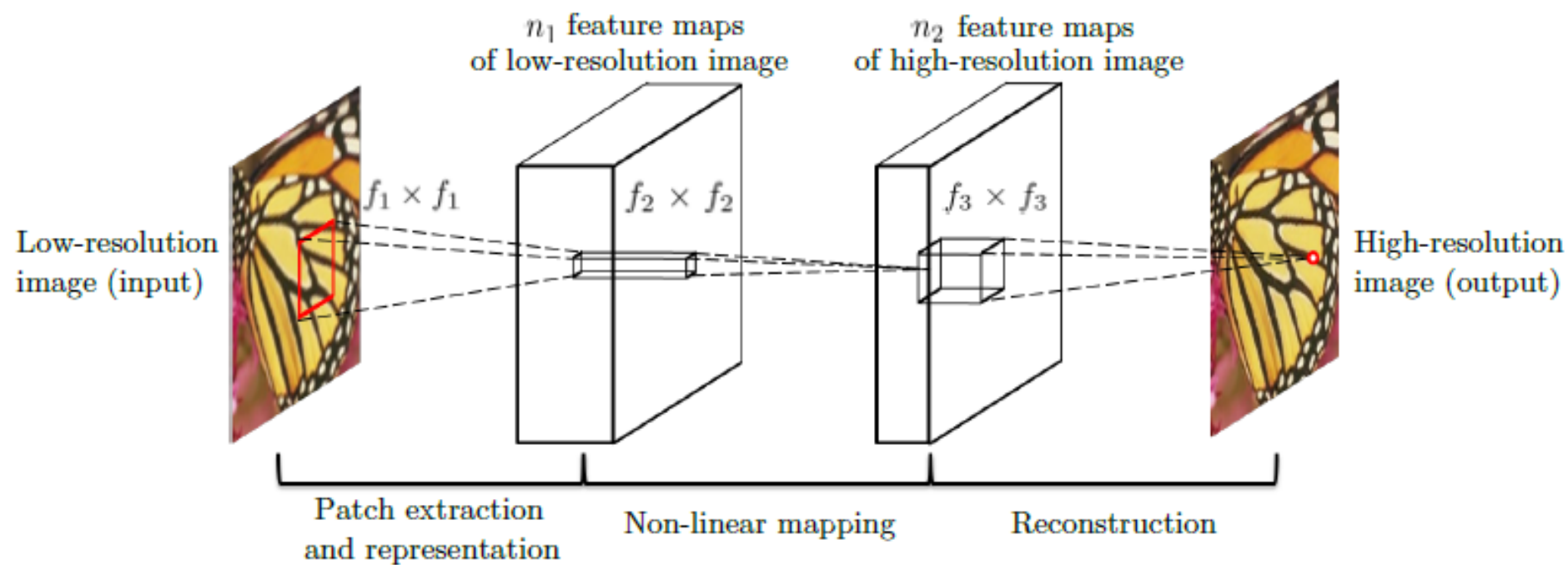
Super-Resolution



고해상도 이미지
(High resolution)

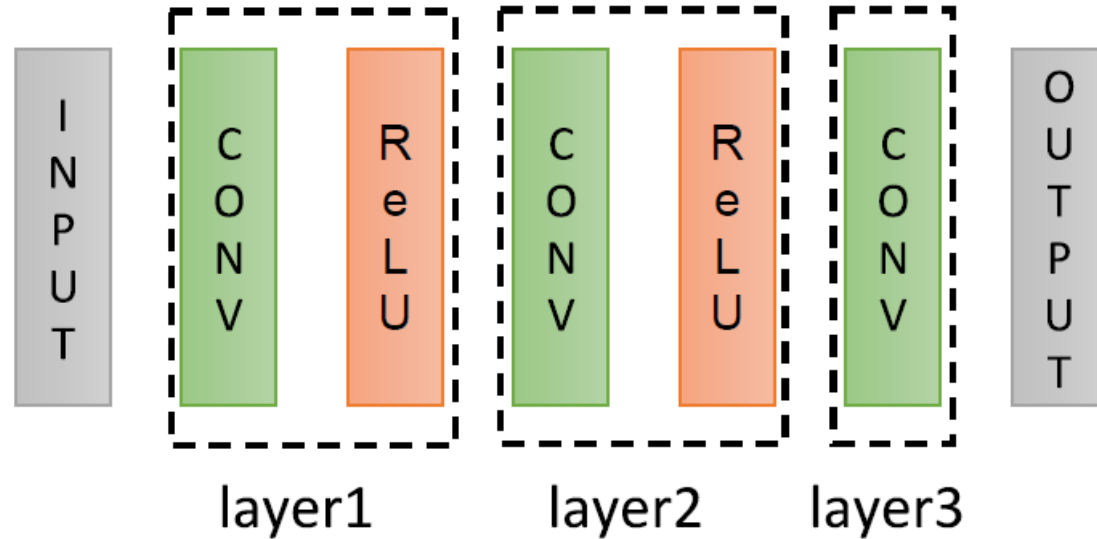
SRCNN

2014 / Chao Dong



SRCNN

2014 / Chao Dong



$$F_1(\mathbf{Y}) = \max(0, W_1 * \mathbf{Y} + B_1)$$

$$F_2(\mathbf{Y}) = \max(0, W_2 * F_1(\mathbf{Y}) + B_2)$$

$$F(\mathbf{Y}) = W_3 * F_2(\mathbf{Y}) + B_3.$$

$$L(\Theta) = \frac{1}{n} \sum_{i=1}^n ||F(\mathbf{Y}_i; \Theta) - \mathbf{X}_i||^2.$$

$$PSNR = 10 \cdot \log_{10} \left(\frac{MAX_I^2}{MSE} \right)$$

SRCNN

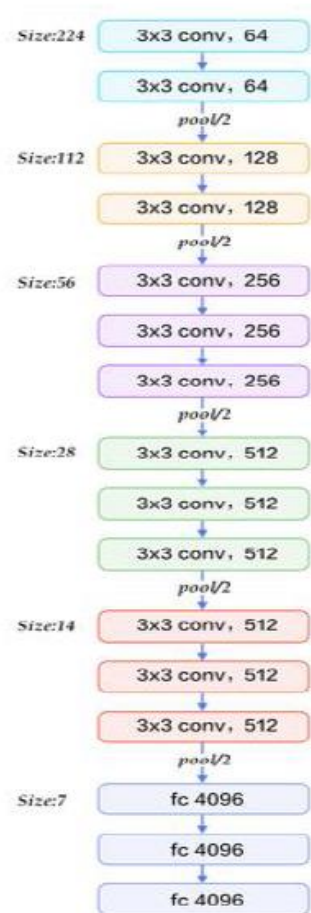
2014 / Chao Dong

- 의의

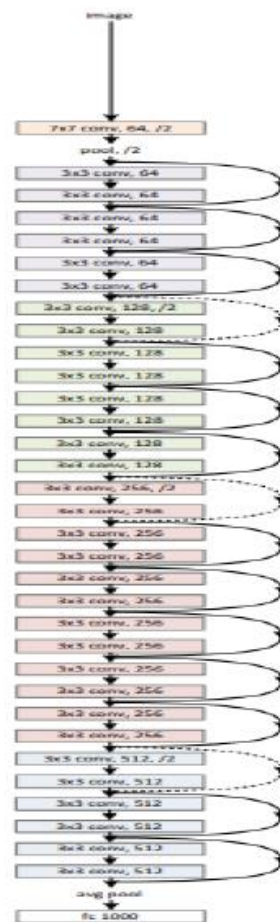
- 그전에 state of art 방법인 sparse coding을 CNN 관점에서 해석
- End-to-end learning으로 그전 방법보다 더 빠르고 좋은 성능 냄
- SR분야에서 최초로 딥러닝 적용
- 그 이후 SR 연구 다 SRCNN 기반으로 함

이후 SR 모델에 영향을 준 모델들

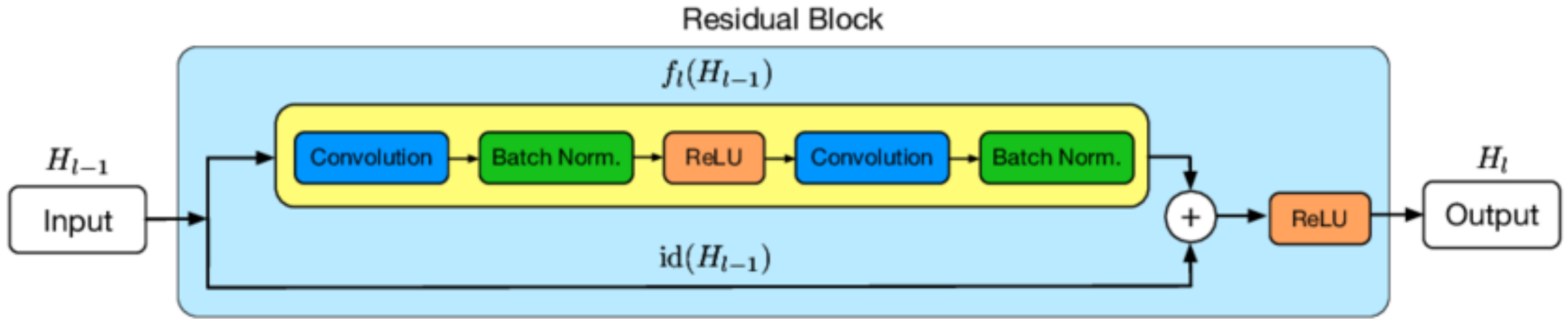
VGG-net



Resnet

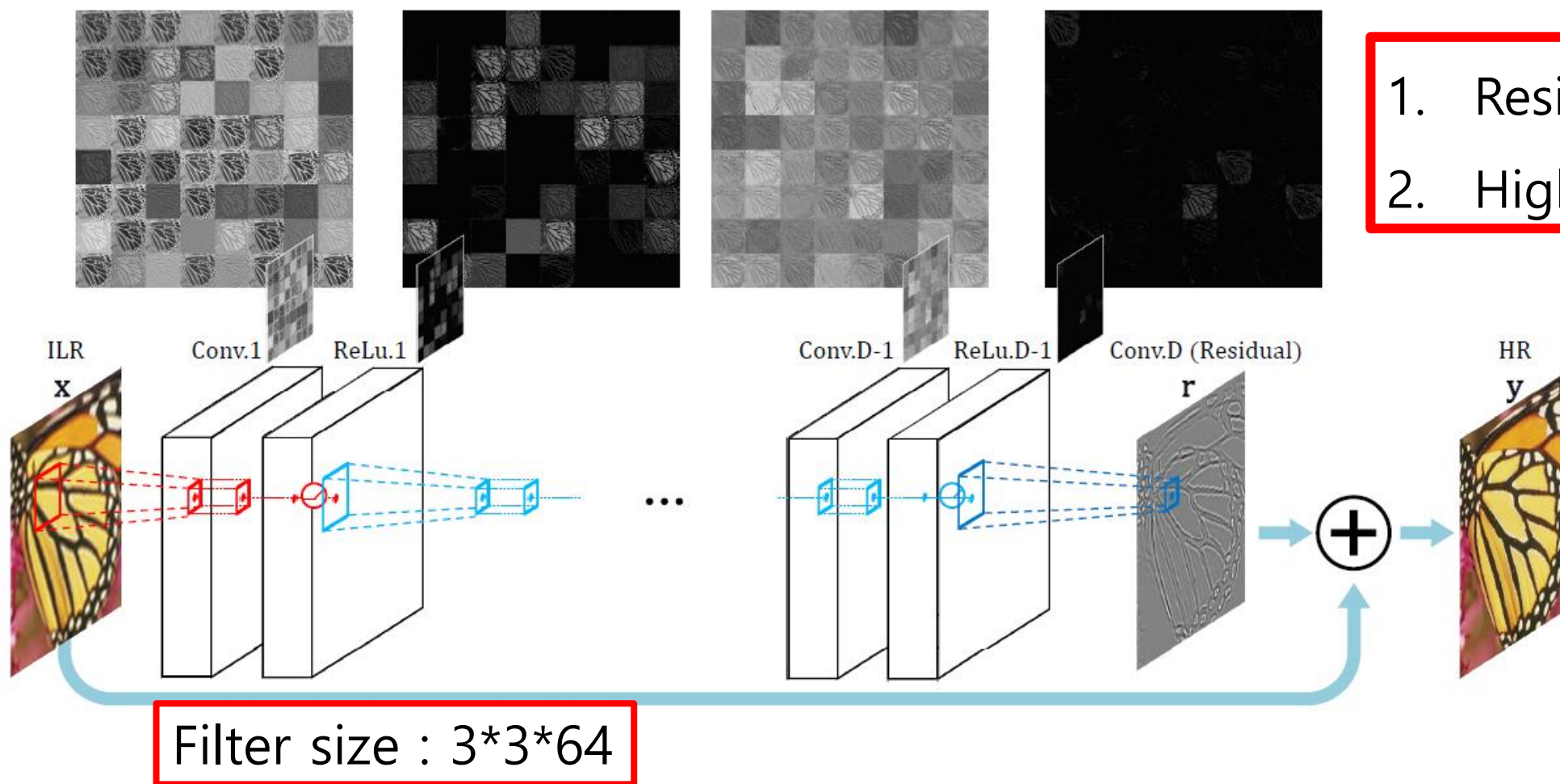


Resnet



VDSR

2015 / Jiwon Kim



VDSR

2015 / Jiwon Kim

- Convolutional Neural Networks with 20 layers
 - Large receptive fields (41 x 41) with consecutive 3x3 convolution kernels
 - Convolutional layers with zero-padding

- Residual learning

- Original loss: $MSE \frac{1}{2} \|y - f(x)\|^2$
- Residual loss: $MSE \frac{1}{2} \|r - f(x)\|^2$, where $r = y - x$.

y: output
x: input
f(x): prediction

- Adjustable gradient clipping with high learning rates

- Common gradient clipping: $[-\theta, \theta]$
- Adjustable gradient clipping: $[-\frac{\theta}{\gamma}, \frac{\theta}{\gamma}]$

θ : predefined range
 γ : current learning rate

VDSR

2015 / Jiwon Kim

- Single CNN for multi-scale factor super-resolution
 - Size of input patch = Size of receptive fields (41 x 41)
 - Images are divided into sub-images with no overlap.
 - Mini-batch size is 64.
 - Sub-images from different scales can be in the same batch.
- Use Zero padding
- Use only Y channel of YCbCr
- Only 80 epoch, 6 hour to train

VDSR

2015 / Jiwon Kim

- Hyperparameters (Final model)
 - **depth** = 20
 - **batch_size** = 64
 - **momentum** = 0.9
 - **weight_decay** = 0.0001
 - **weight_initialization** = he_normal()
 - **activations** = ReLU
 - **epochs** = 80
 - **learning_rate** = 0.1
 - **learning_rate** was decreased by a factor 10 every 20 epochs.

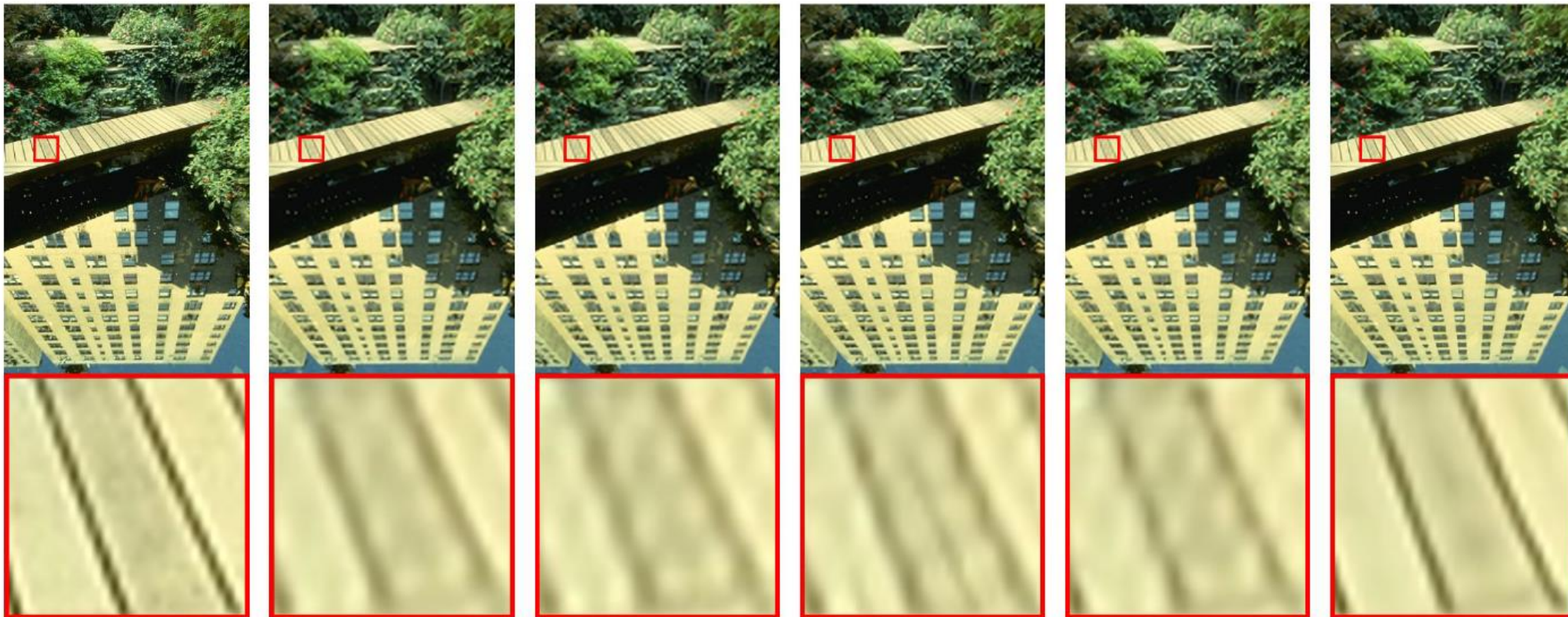
VDSR

2015 / Jiwon Kim

	SRCNN	VDSR
Architecture	end-to-end learning	end-to-end learning
Receptive fields	13 x 13	41 x 41
Scale factors	x3	x2, x3, x4
Learning rate	10^{-5}	1/10 learning rate decay (10^{-2} , 10^{-6})
Depth	4	(Up to) 20

VDSR

2015 / Jiwon Kim



Ground Truth
(PSNR, SSIM)

A+ [22]
(22.92, 0.7379)

RFL [18]
(22.90, 0.7332)

SelfEx [11]
(23.00, 0.7439)

SRCNN [5]
(23.15, 0.7487)

VDSR (Ours)
(23.50, 0.7777)

VDSR

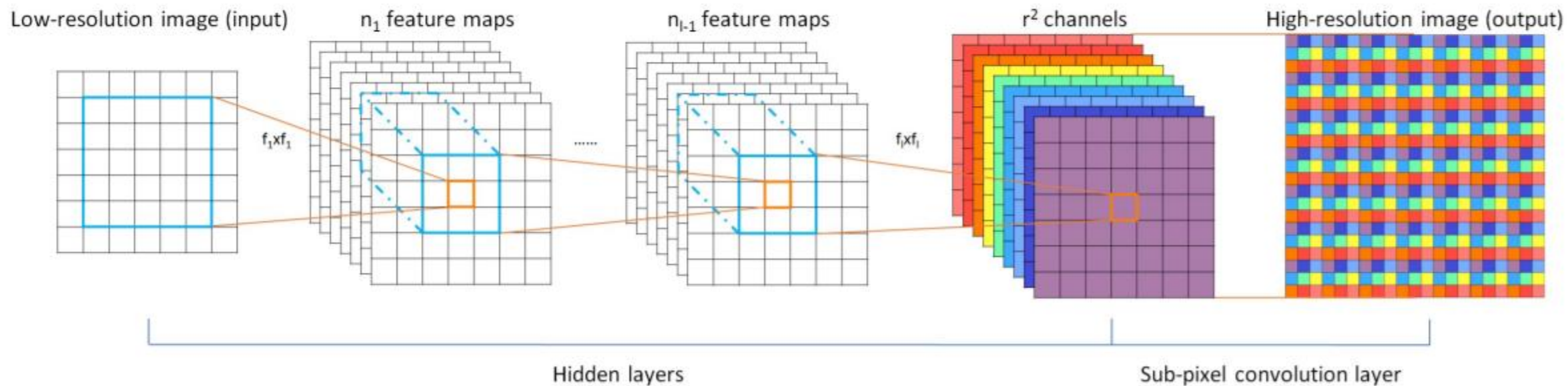
2015 / Jiwon Kim

- 의의

- SRCNN의 레이어를 20까지 확장
- Residual learning, adjustable gradient clipping 통해 빠른 학습
- Residual 구조의 가능성 보여줌

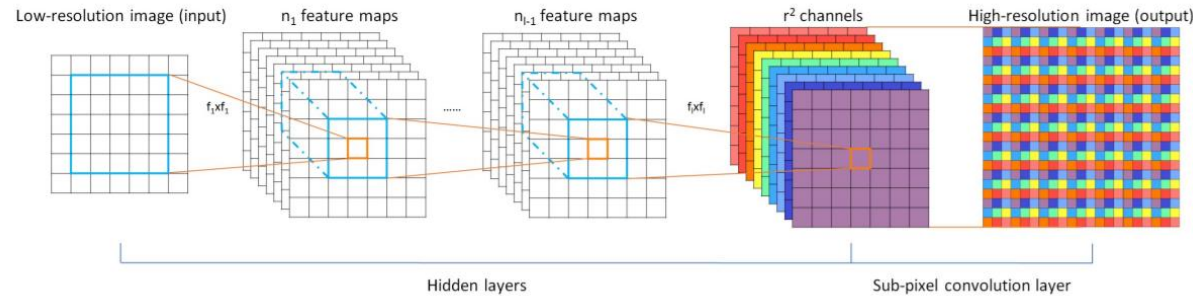
ESPCN

2016 / Wenzhe Shi



ESPCN

2016 / Wenzhe Shi

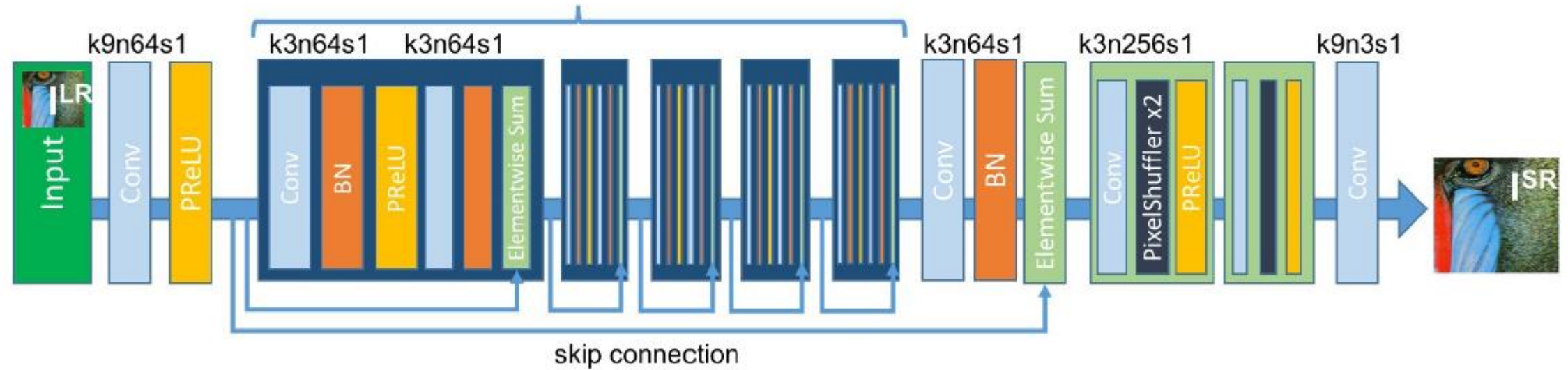


- 그전에 bicubic interpolation 으로 LR HR로 늘리고 시작
 - 이게 비효율적이고 학습을 더 어렵게 만든다고 생각함
- Network 단계에서 upscale
- Realtime 적용이 가능해 동영상에도 적용 가능
- 이후 SR network 다 이 방법 적용

SRResnet

2017 / Christian Ledig

Generator Network

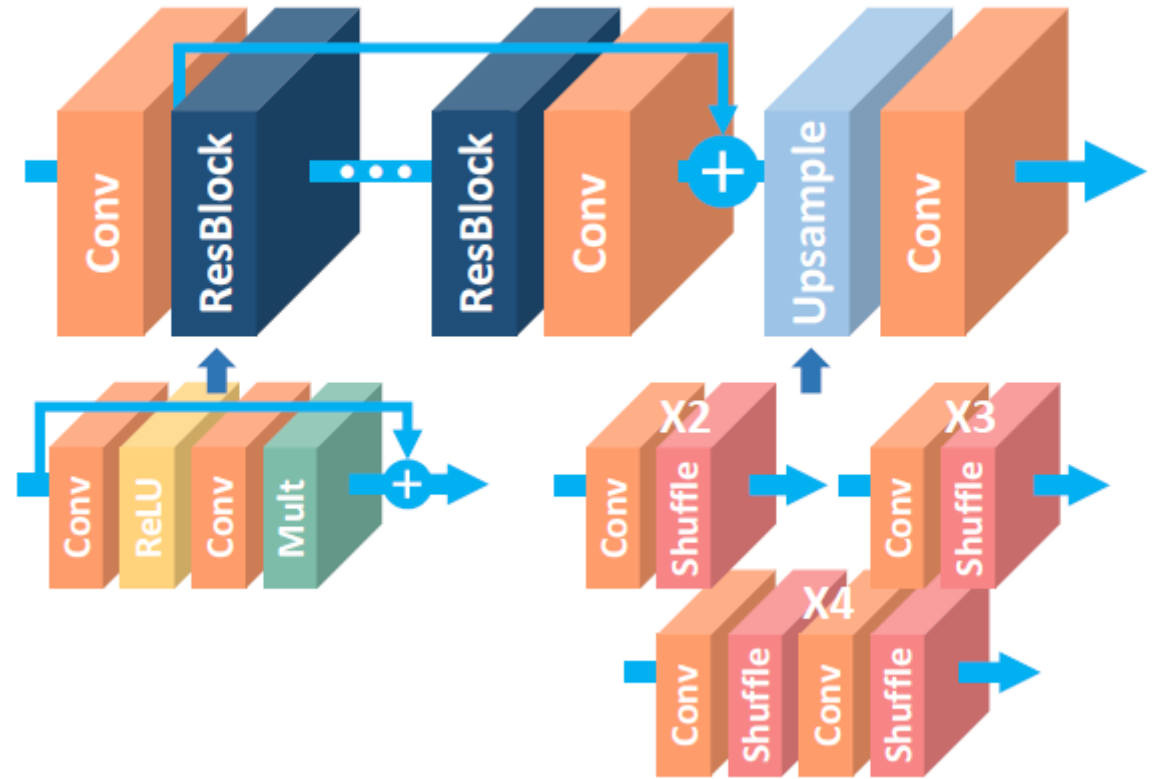


EDSR

2017 / Bee Lim

- 특징

- 불필요한 BN, Relu등을 제거하고 더 깊은 Residual block 쌓아서 성능 높임
- L2 loss가 아닌 L1 loss 사용
 - (둘다 해봤더니 L1이 더 잘나와서)



RDN

2018 / Yulun Zhang

- Just Residual model
 - VDSR 모델은 하나의 Skip-connection 만 존재
 - SRresnet 모델은 Residual block 과 local skip-connection 적용
 - EDSR 모델은 SRresnet 구조에서 발전된 모델
- 문제점:
 - 각 convolutional layer 의 정보들을 충분히 사용하지 못하고 무시해버린다.

Dense Net

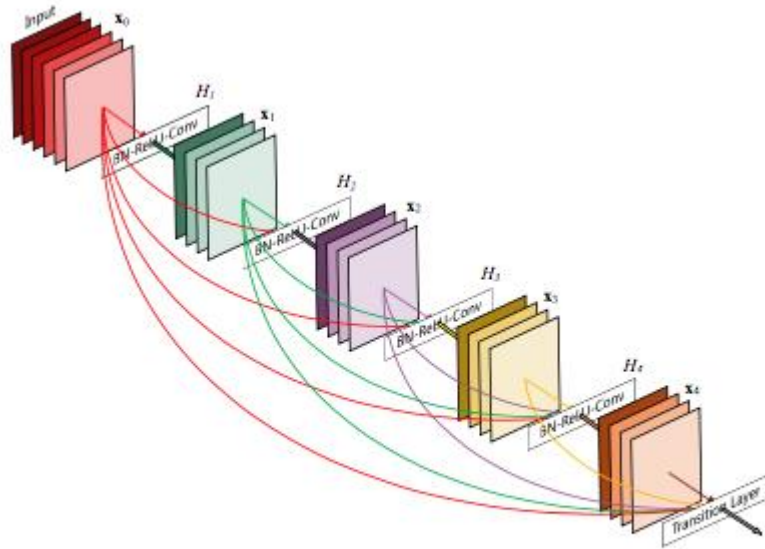


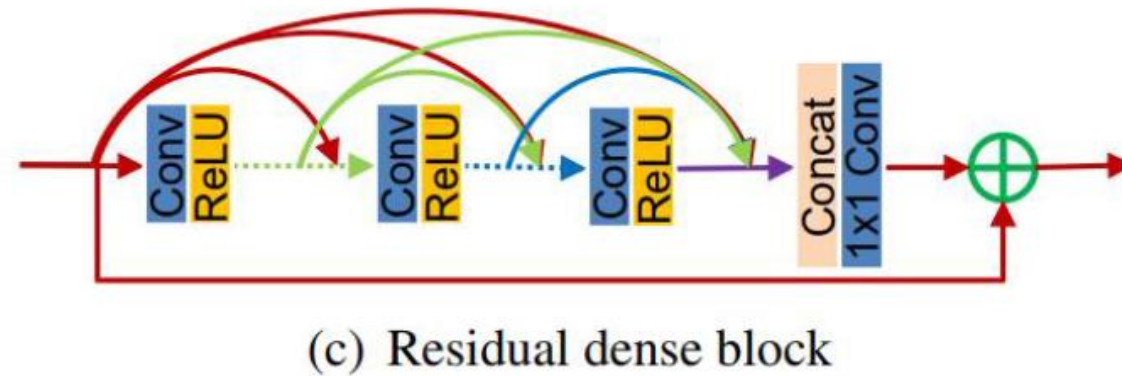
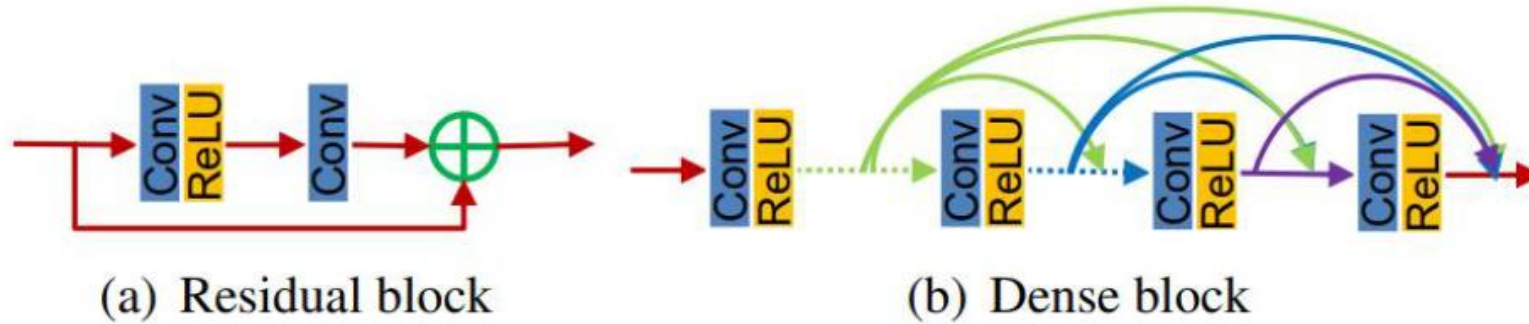
Figure 1: A 5-layer dense block with a growth rate of $k = 4$. Each layer takes all preceding feature-maps as input.

- 모든 레이어는 서로 연결
 - 단 한방향으로만
- $L(L+1)/2$ 만큼의 새로운 커넥션이 생김
- Resnet에 비해 매우 Narrow
- Resnet 방식 : $x_l = H_l(x_{l-1}) + x_{l-1}$
- Dense net 방식 : $x_l = H_l([x_0, x_1, \dots, x_{l-1}])$

RDN

2018 / Yulun Zhang

Residual + Dense



RDN

2018 / Yulun Zhang

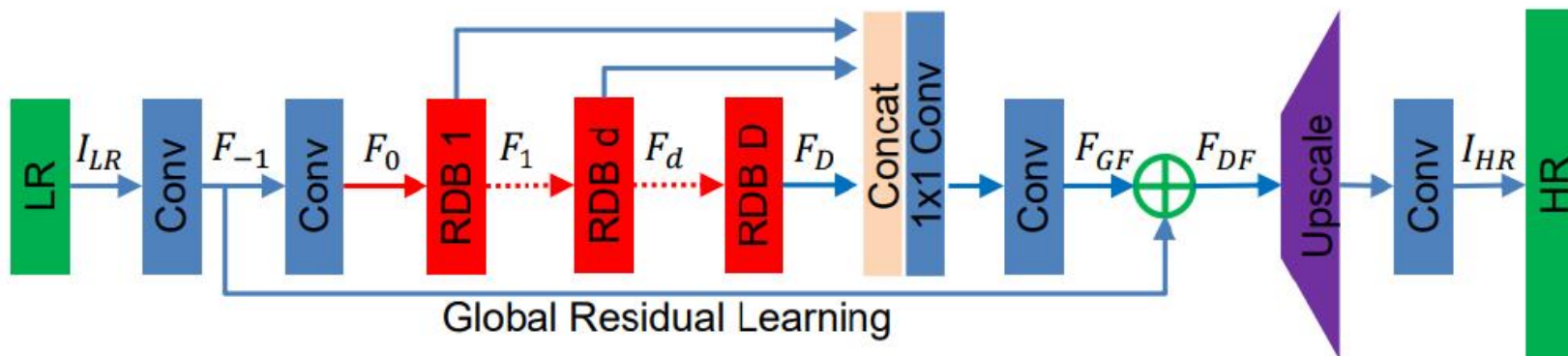


Figure 2. The architecture of our proposed residual dense network (RDN).

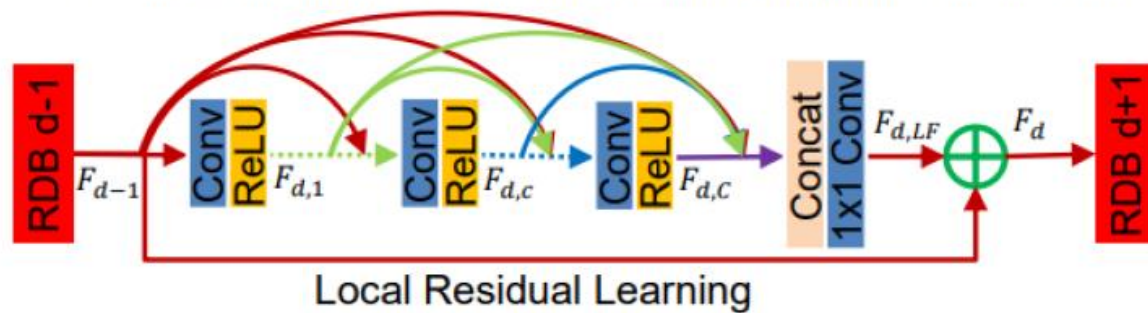


Figure 3. Residual dense block (RDB) architecture.

- 선행 Residual block(RDB)의 정보를 현재 RDB에서 직접 접근
- 선행 RDN의 features을 Global Feature Fusion 으로 합침
- Efficient sub-pixel(ESPCN) 기법으로 upscaling

RDN

2018 / Yulun Zhang

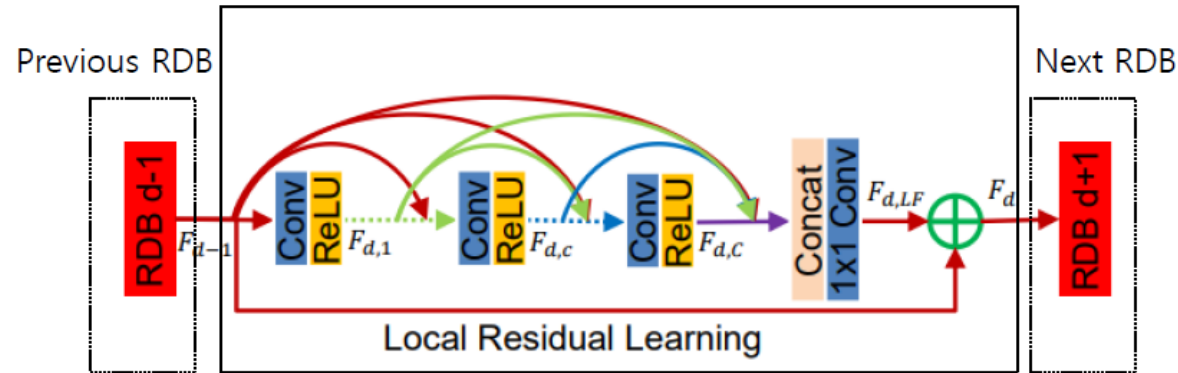


Figure 3. Residual dense block (RDB) architecture.

- Contiguous memory : 선행의 RDB 정보가 현재의 RDB 내부로 직접 전달
- Dense connection : 선행 local layers의 features를 직접 전달
- Local Feature Fusion : RDN 안의 local layers 의 features 를 합치는 과정 (1*1 conv 말함)

SRGAN

2017 / Christian Ledig

- PSNR은 여전히 유효한 지표인가?

bicubic
(21.59dB/0.6423)



SRResNet
(23.53dB/0.7832)



SRGAN
(21.15dB/0.6868)



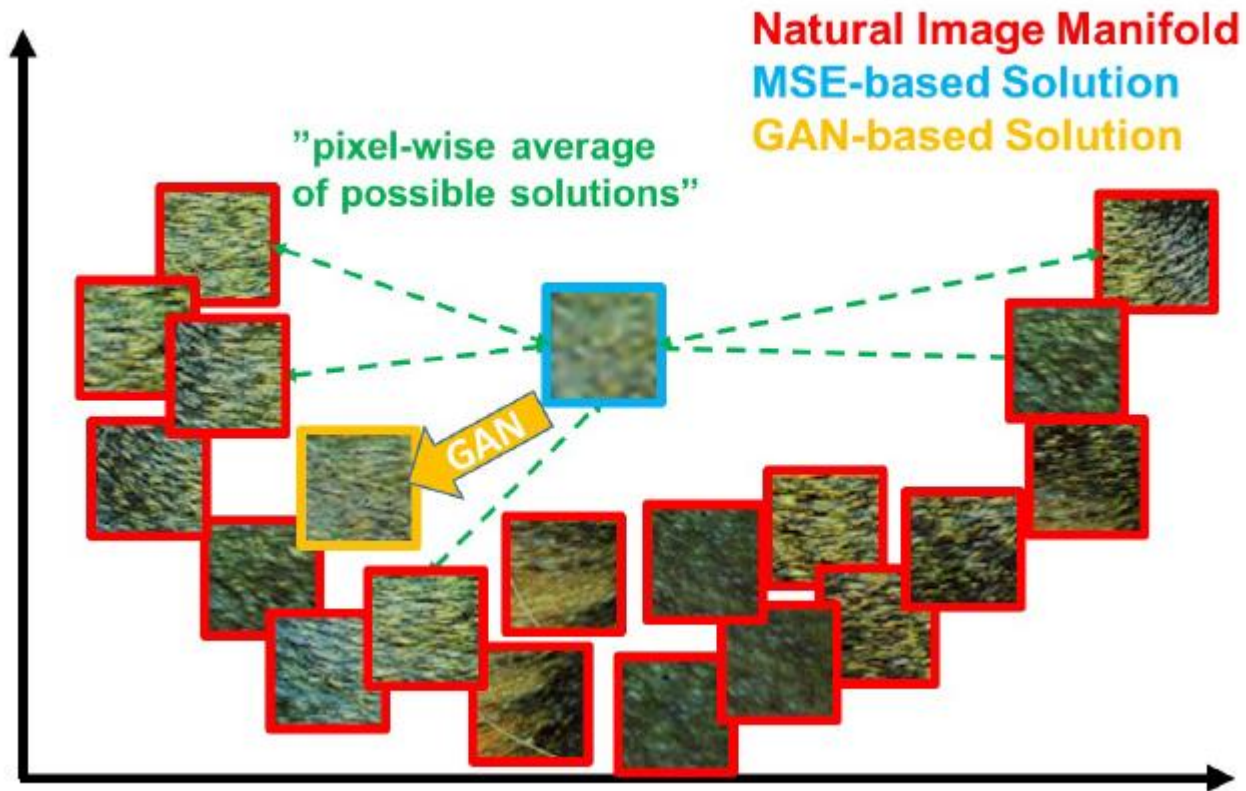
original



SRGAN

2017 / Christian Ledig

- PSNR은 여전히 유효한 지표인가?



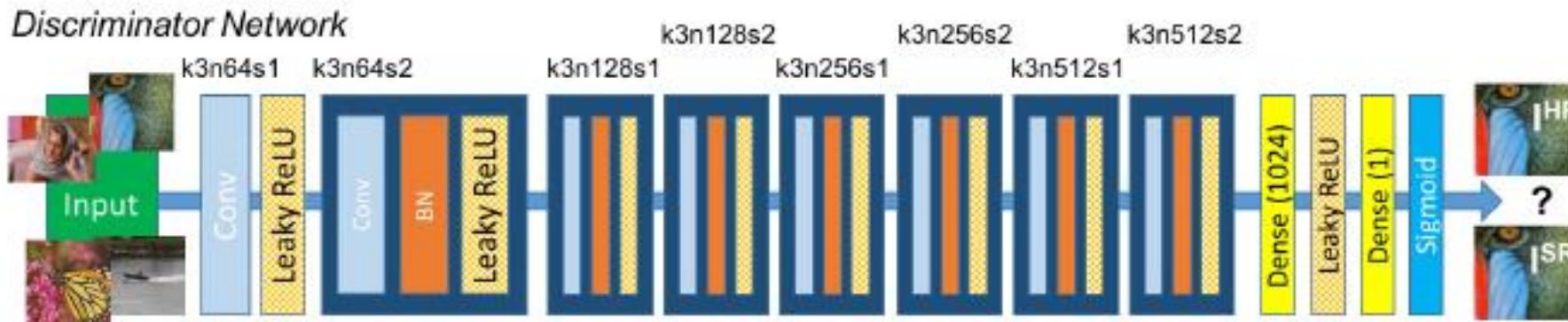
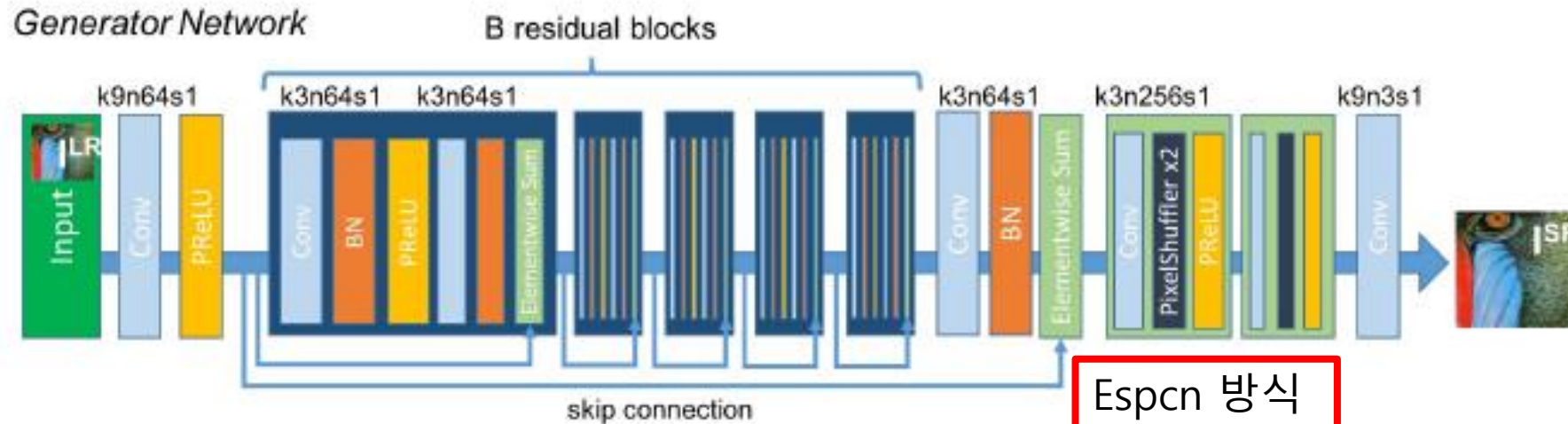
MSE optimize 하서는 high frequency feature 놓치게 됨!

SRGAN

2017 / Christian Ledig

ParametricReLU

No pooling



Leaky ReLU

SRGAN

2017 / Christian Ledig

$$l_{MSE}^{SR} = \frac{1}{r^2WH} \sum_{x=1}^{rW} \sum_{y=1}^{rH} (I_{x,y}^{HR} - G_{\theta_G}(I^{LR})_{x,y})^2$$

$$l^{SR} = \underbrace{l_X^{SR}}_{\text{content loss}} + \underbrace{10^{-3}l_{Gen}^{SR}}_{\text{adversarial loss}}$$

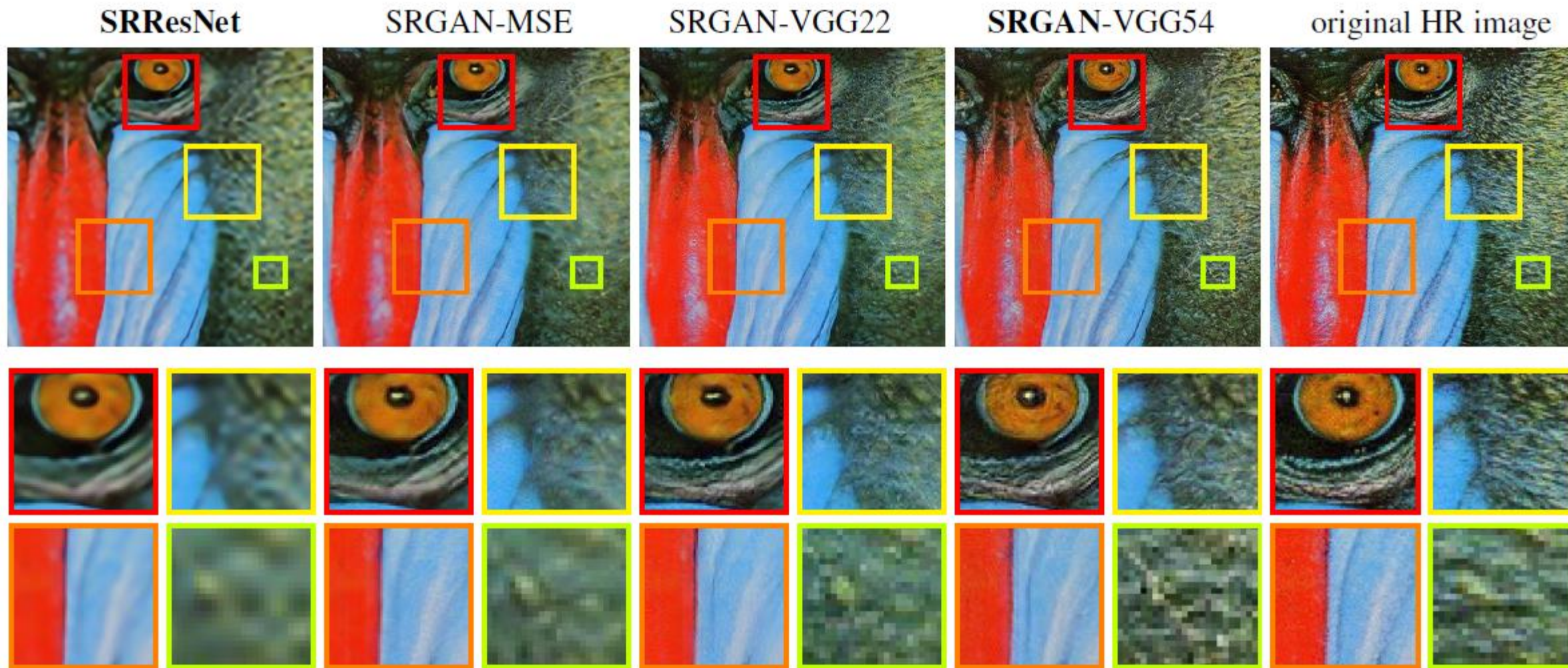
perceptual loss (for VGG based content losses)

$$l_{VGG/i,j}^{SR} = \frac{1}{W_{i,j}H_{i,j}} \sum_{x=1}^{W_{i,j}} \sum_{y=1}^{H_{i,j}} (\phi_{i,j}(I^{HR})_{x,y} - \phi_{i,j}(G_{\theta_G}(I^{LR}))_{x,y})^2$$

$$l_{Gen}^{SR} = \sum_{n=1}^N -\log D_{\theta_D}(G_{\theta_G}(I^{LR}))$$

SRGAN

2017 / Christian Ledig



정리

- 2014 : Deep learning Super Resolution의 시작 – SRCNN
- 2015 : VGG-net 과 Resnet의 등장 후 SR 영역으로 확대 – VDSR
- 2016 : Residual 구조의 여러 방향과 Upscaling의 관점 제시 – ESPCN, VDSR
- 2017 : Resnet 구조의 적용과 발전, 그리고 GAN을 이용한 SR 모델 등장
– SRresnet, ESDR, SRGAN
- 2018(진행중) : Residual 과 Dense의 결합을 통한 새로운 Residual SR 모델 제시 - RDN

References

- https://www.youtube.com/watch?v=1jGr_OFyfa0&t=653s
- <https://www.youtube.com/watch?v=nvsYKSHw0jo&t=3051s>