

---

# Introduction to Image Super-Resolution

---

발표자 : 백인성

2021.04.09.

# 발표자 소개

---

Insung Baek



## ❖ 백인성 (Insung Baek)

- Korea University
- Data Mining & Quality Analytics Lab
- Ph.D. Student (2018. 9 ~ Present)

## ❖ Research Interest

- Explainable Artificial Intelligence Algorithms
- Game Artificial Intelligence
- Image Super-Resolution

## ❖ Contact

- E-mail: [insung\\_baek01@korea.ac.kr](mailto:insung_baek01@korea.ac.kr)

# 목차

---

1. Introduction
2. Image Super-Resolution
3. Image Super-Resolution using Deep Convolutional Networks (SRCNN)
4. Second-order Attention Network for Single Image Super-Resolution (SAN)
5. Conclusion

---

# 1. Introduction

---

# Introduction

---

Image Super-Resolution 약어

- ✓ **SR** = Super-Resolution (초고해상도)  
→ 세미나에서는 주로 모델로 생성된 Image를 의미
- ✓ **HR** = High-Resolution (고해상도)  
→ 세미나에서는 주로 Output Image를 의미
- ✓ **LR** = Low-Resolution (저해상도)  
→ 세미나에서는 주로 Input Image를 의미

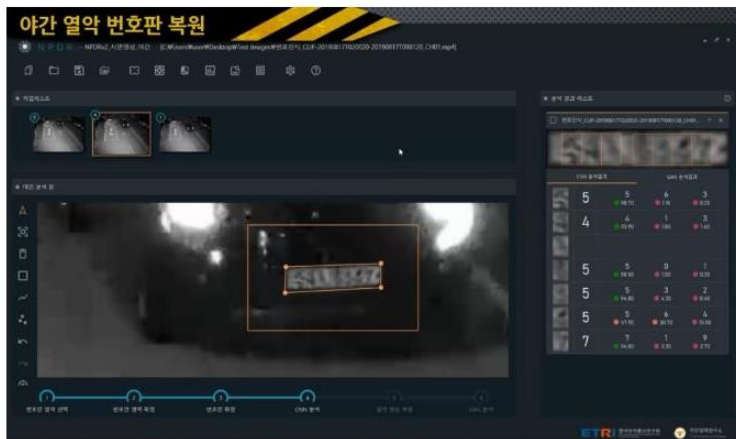
# Introduction

## Image Super-Resolution 정의

- ❖ Image Super-Resolution은 저해상도 이미지를 고해상도로 변환시키는 문제를 의미함
- ❖ 미디어, 의료, 보안 등 다양한 산업 분야에서 중요한 문제로 대두되고 있음

### CCTV에 찍힌 흐릿한 차량 번호 10분 만에 알아낸다

| ETRI "해상도 높여 번호판 복원...인공지능 'NPDR' 기술 개발"



<ETRI - 차량 번호판 복원>

### 렘브란트 대표작 '야경' AI 통해 디지털 복원

17세기 명화, 45억만화소 디지털 이미지로 재탄생  
디지털 스캔과 인공지능 기술 활용, 고미술품 복원에 새 지평 열어



디지털로 복원된 렘브란트의 대표작 '야경' /사진=hyper-resolution.org

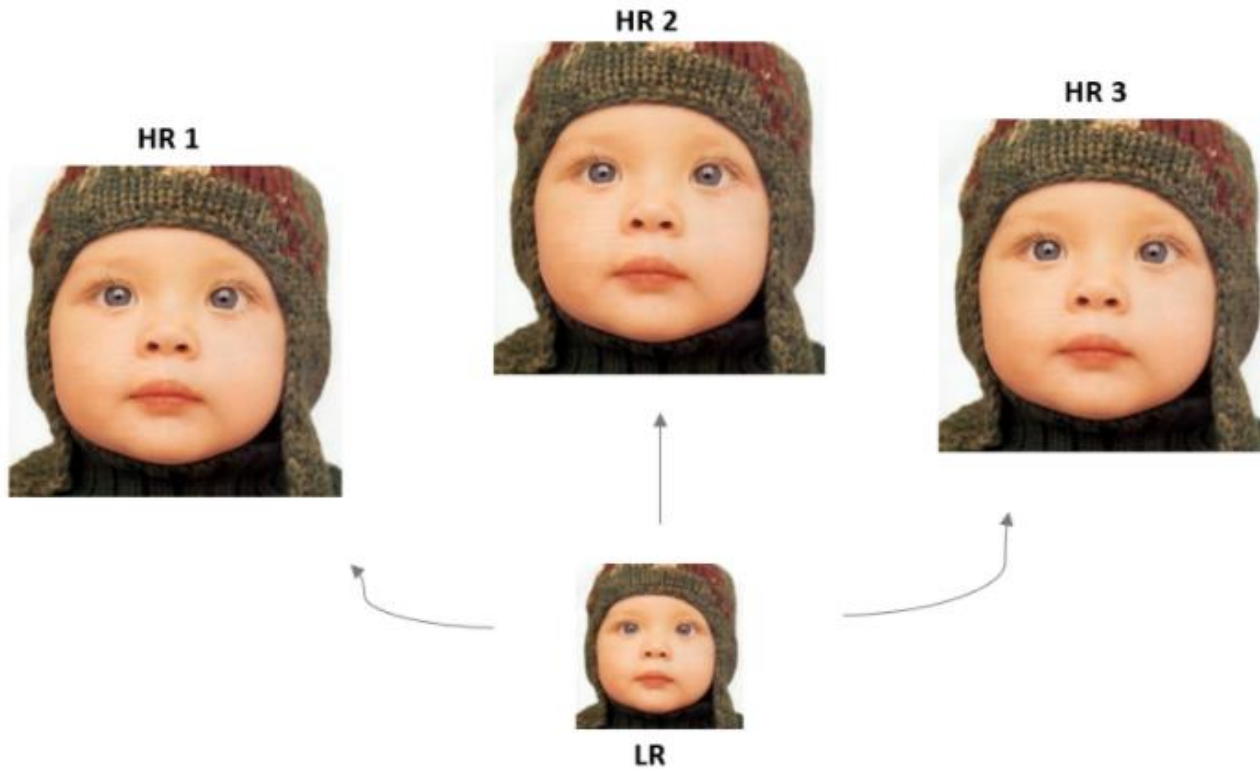
<암스테르담 박물관-미술 명작 복원>

출처: <https://www.yna.co.kr/view/AKR20191108058400063>  
<http://www.digitaltoday.co.kr/news/articleView.html?idxno=233998>

# Introduction

Image Super-Resolution 문제의 어려움

- ❖ 하지만 Image Super-Resolution는 유일한 정답이 존재하지 않는 ill-posed problem 존재
- ❖ 저해상도 이미지를 고해상도로 복원할 때 정답이 단 1개만 존재하지 않음



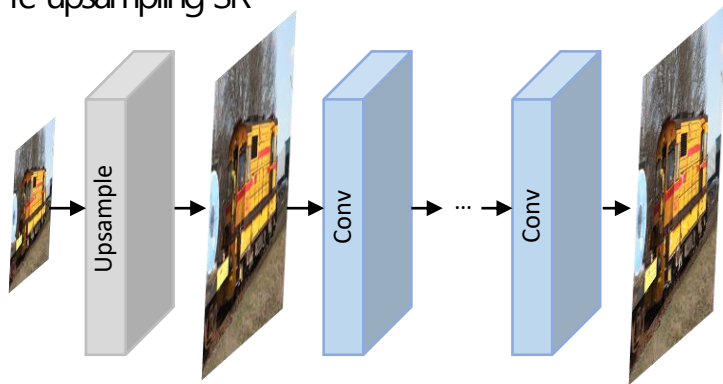
<ill-posed problem example>

# Introduction

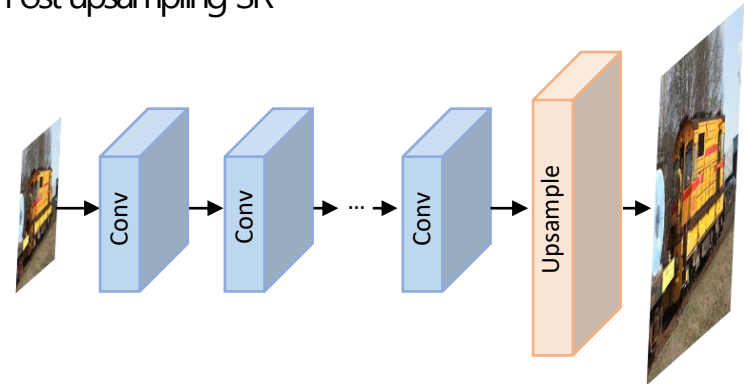
딥러닝 모델을 적용한 Image Super-Resolution

- ❖ Computer Vision 영역에서 좋은 성능을 보이고 있는 다양한 딥러닝 모델을 적용해 Image Super-Resolution 문제를 해결하려는 연구가 다수 진행되고 있음

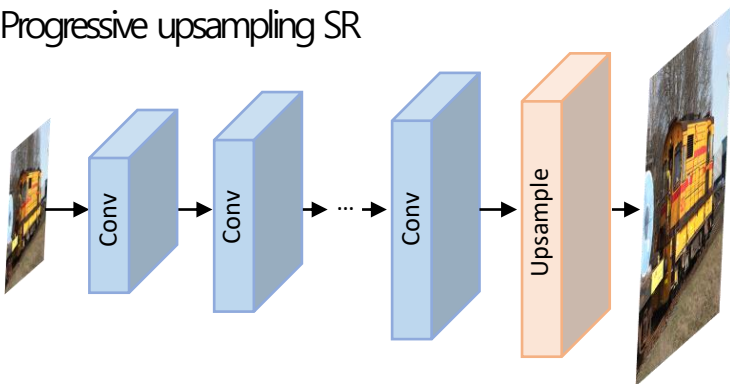
① Pre-upsampling SR



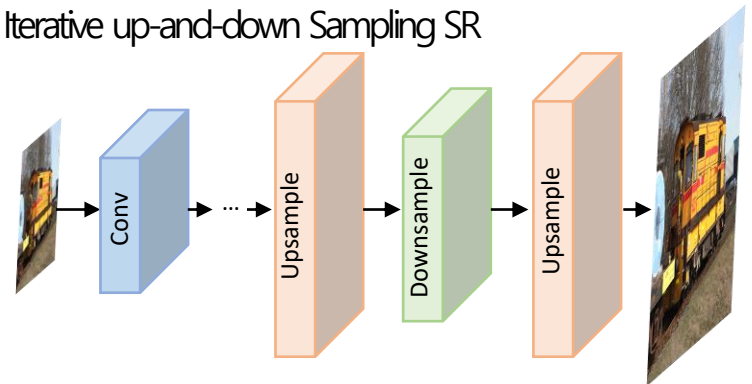
② Post-upsampling SR



③ Progressive upsampling SR



④ Iterative up-and-down Sampling SR



<Deep Learning Frameworks for Super-Resolution>

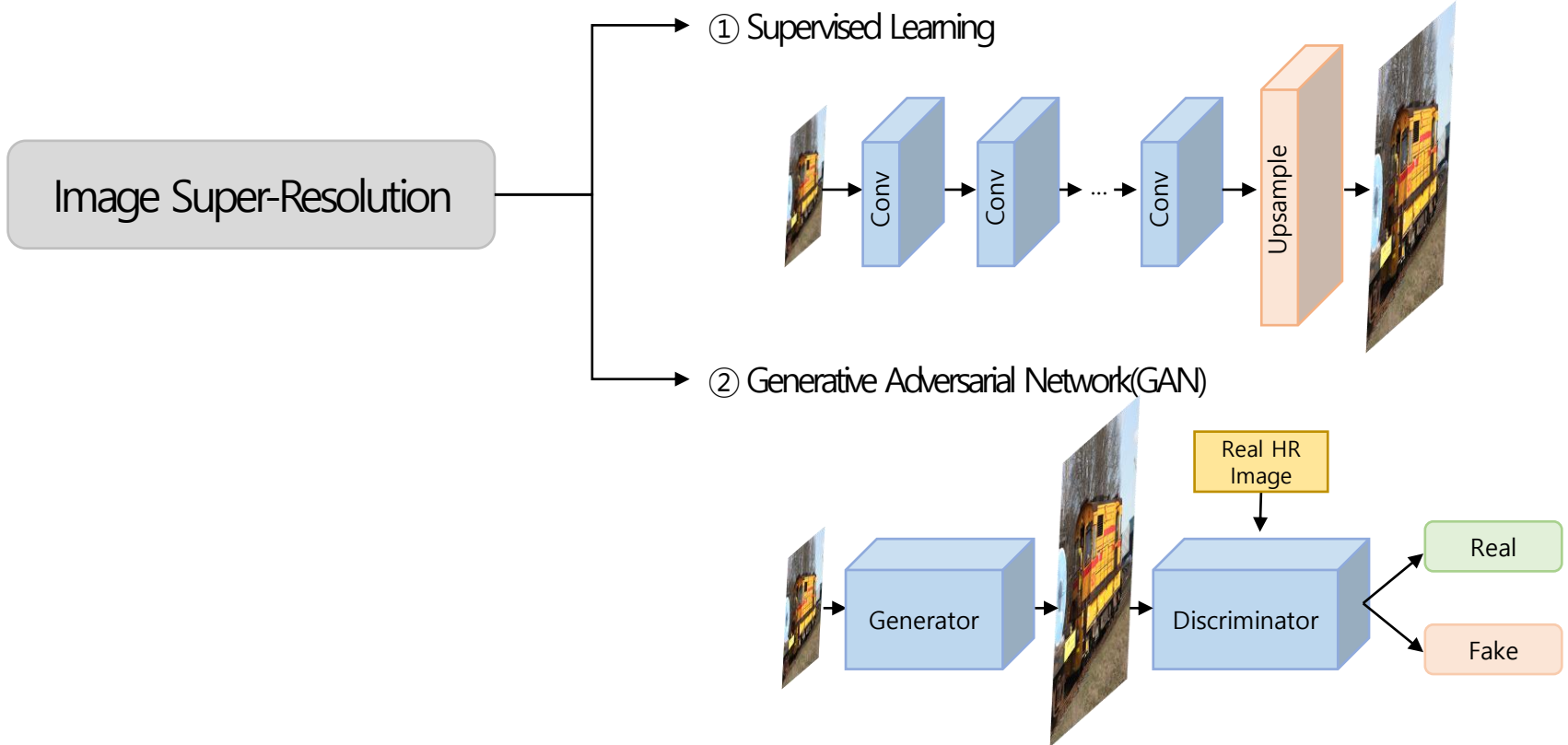
출처: Wang, Z., Chen, J., & Hoi, S. C. (2020). Deep learning for image super-resolution: A survey. *IEEE transactions on pattern analysis and machine intelligence*.



# Introduction

## Image Super-Resolution 분류

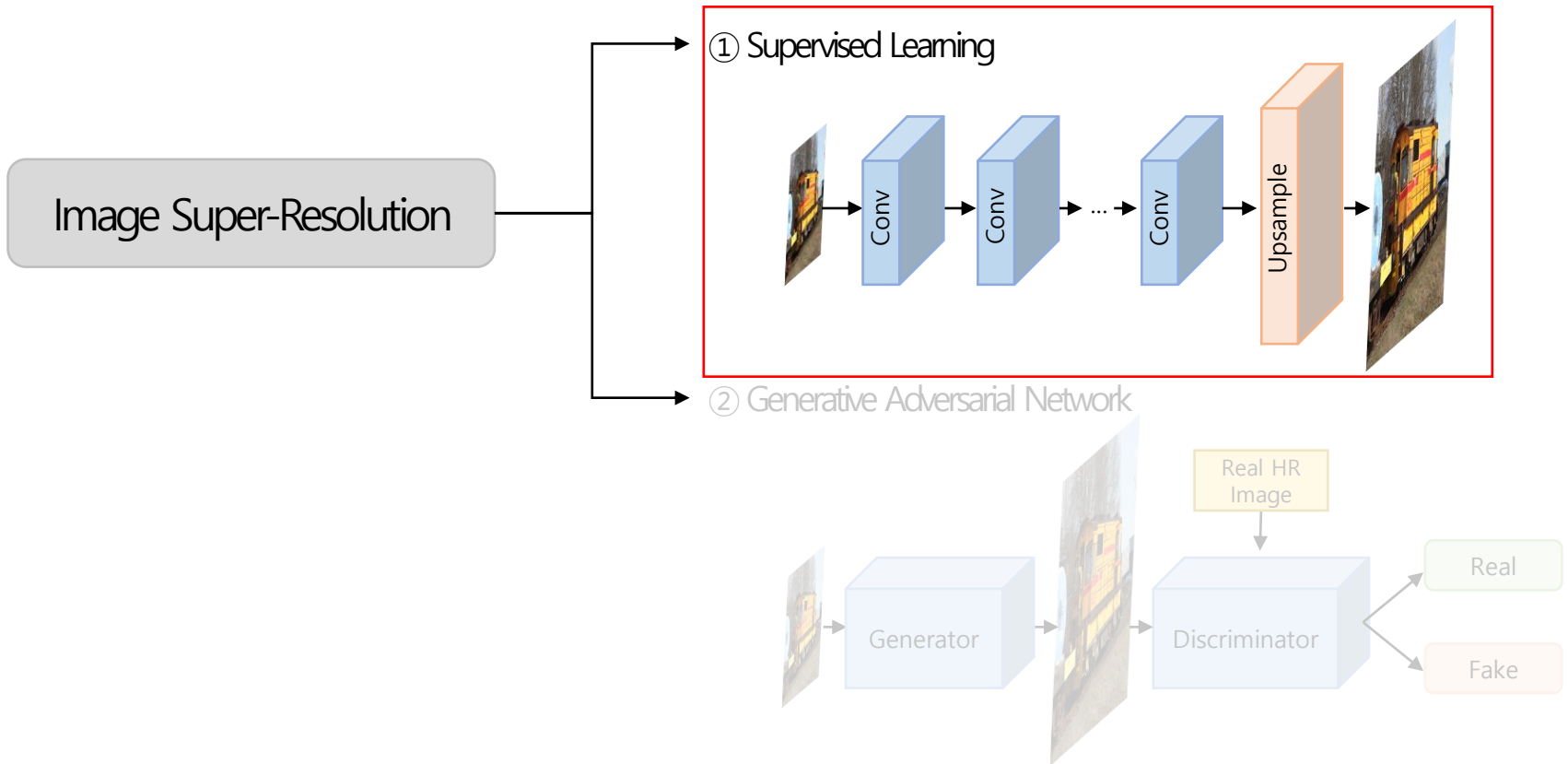
- ❖ 저해상도 이미지로 고해상도 이미지를 예측하는 Supervised Learning 방식
- ❖ Generative Adversarial Network(GAN)을 활용해 고해상도 이미지를 생성하는 방식



# Introduction

## Image Super-Resolution 분류

- ❖ 저해상도 이미지로 고해상도 이미지를 예측하는 Supervised Learning 방식
- ❖ GAN(Generative Adversarial Network)을 활용해 고해상도 이미지를 생성하는 방식
- ❖ 본 세미나에서는 Supervised Learning 방식을 중심으로 설명이 진행될 예정



# Introduction

---

## Image Super-Resolution 특징

1. 저해상도 이미지와 고해상도 이미지 크기가 다른 문제를 어떻게 다룰 것인가?
2. Image Super-Resolution 문제에 어떻게 딥러닝 모델을 적용할 것인가?
3. 고해상도 이미지 변환이 잘 되었다는 것을 어떻게 평가할 것인가?

---

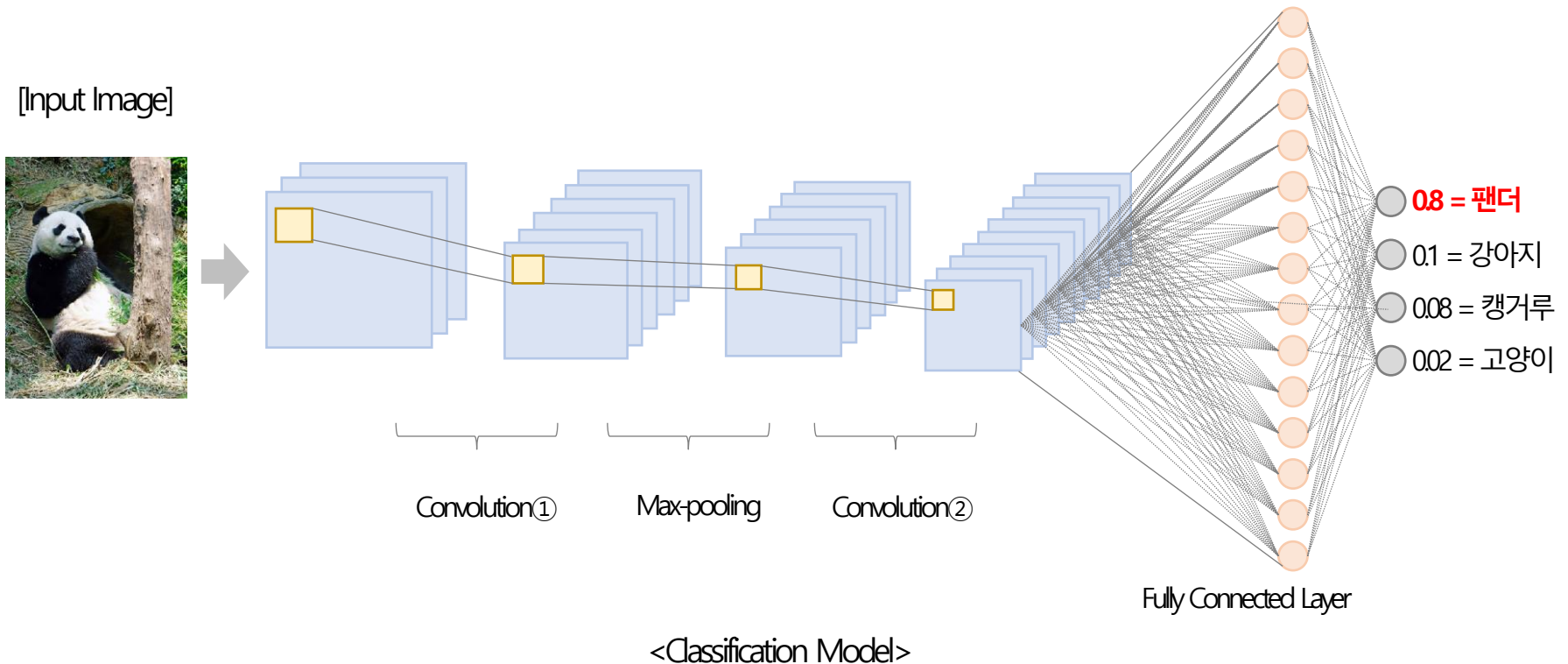
## 2. Image Super-Resolution

---

# Image Super-Resolution

이미지 분류 문제에서 일반적인 딥러닝 모델

- ❖ 최근 Computer vision 문제를 해결하기 위한 다양한 딥러닝 모델이 연구되고 있음
- ❖ 일반적인 이미지 분류 문제에서는 Convolutional Layer와 Fully Connected Layer를 사용해 딥러닝 모델을 구축하게 됨

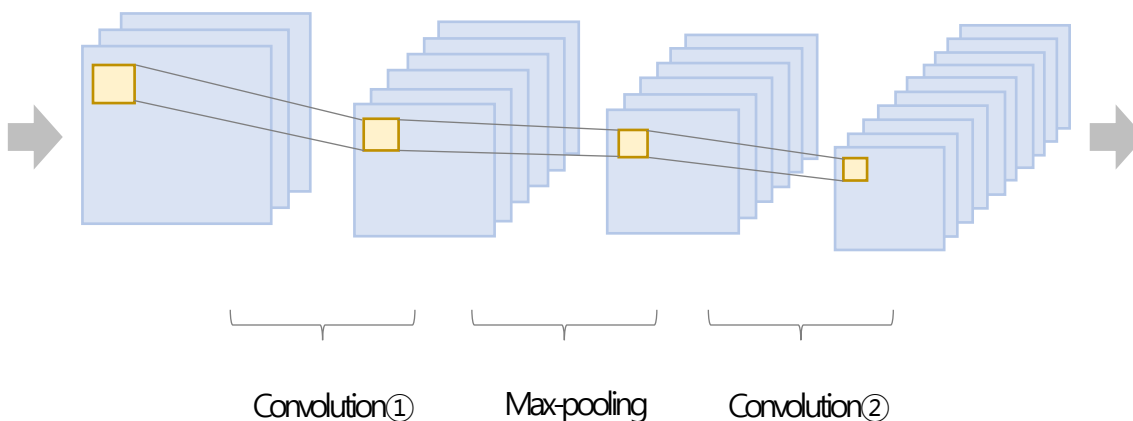


# Image Super-Resolution

Image Super-Resolution 문제에서 딥러닝 모델

- ❖ 하지만 Image Super-Resolution 문제를 해결하기 위해서는 추가적인 방법이 필요함

[Input Image]



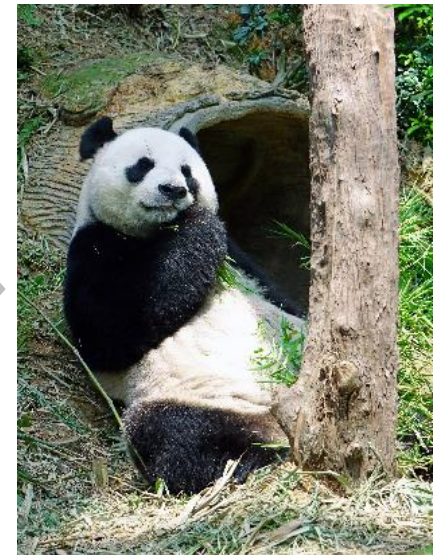
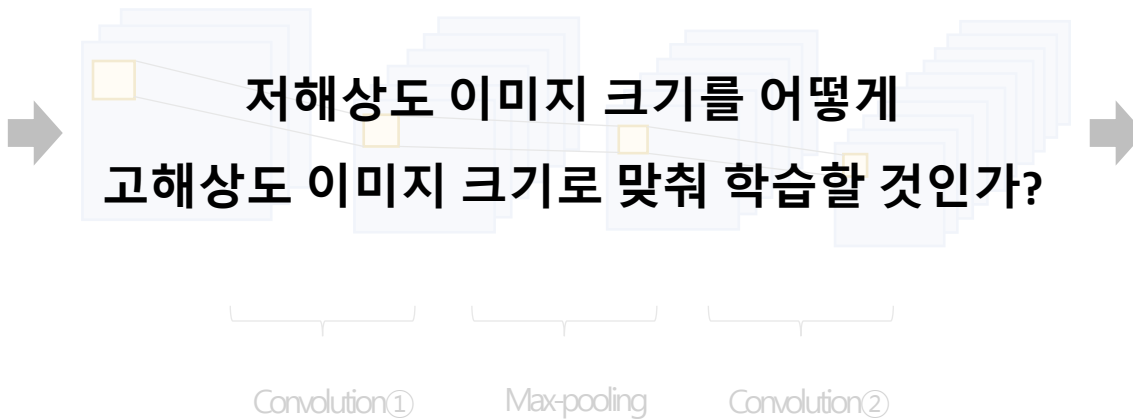
<Image Super-Resolution Model>

# Image Super-Resolution

Image Super-Resolution 문제에서 딥러닝 모델

- ❖ 하지만 Image Super-Resolution 문제를 해결하기 위해서는 추가적인 방법이 필요함
- ❖ 저해상도 이미지 크기와 고해상도 이미지 크기를 맞추기 위한 방법들이 필요함

[Input Image]



<Image Super-Resolution Model>

# Image Super-Resolution

저해상도 이미지와 고해상도 이미지 크기 차이

- ❖ 하지만 Image Super-Resolution 문제를 해결하기 위해서는 추가적인 방법이 필요함
- ❖ 저해상도 이미지 크기와 고해상도 이미지 크기를 맞추기 위한 방법들이 필요함

[Input Image-LR]



1	0
2	3

??

1	1	0	3
0	2	1	0
2	2	3	2
1	0	0	1

[Output Image-HR]





# Image Super-Resolution

저해상도 이미지와 고해상도 이미지 크기 차이

- ❖ 먼저 저해상도인 Input 이미지를 고화질 이미지 크기를 맞춰야 함
- ❖ 크기를 맞춘 뒤, 빈 영역에 대한 값을 어떻게 채울 것인가에 대한 고민이 필요

[Input Image-LR]



1	0
2	3

Resizing

1			0
2			3

빈 곳을 어떻게  
채울 것인가?

[Output Data]

1	1	3	0
0	2	1	0
1	2	2	0
2	0	0	3

\*\*쉽게 보여주기 위한 단순 example임

# Image Super-Resolution

## Upsampling 방법

- ❖ 인접 픽셀 값을 활용해서 빈 영역을 채우는 interpolation 방법
- ❖ Upsampling이 가능한 Layer를 활용해서 upsampling parameter를 학습하는 방법

### ① Interpolation-based upsampling

- Nearest-neighbor Interpolation
  - ✓ 인접 픽셀 값으로 채움
- Bilinear Interpolation
  - ✓ 인접 4개 픽셀 값의 liner interpolation 진행
- Bicubic Interpolation
  - ✓ 인접 16개 픽셀 값의 cubic interpolation 진행

### ② Learning-based upsampling

- Transpose Convolution Layer
  - ✓ 저해상도 이미지 확장 후 Convolution 연산 진행
- Sub-pixel Convolution Layer
  - ✓ Convolution 연산 후 pixel별로 모음

# Image Super-Resolution

## Nearest-neighbor Interpolation

- ❖ Interpolation 방법 중에서 가장 간단한 방법
- ❖ 빈 픽셀은 가장 가까운 픽셀 값으로 채우는 방법

[Input Image-LR]



1	0
2	3

Resizing

1			0
2			3

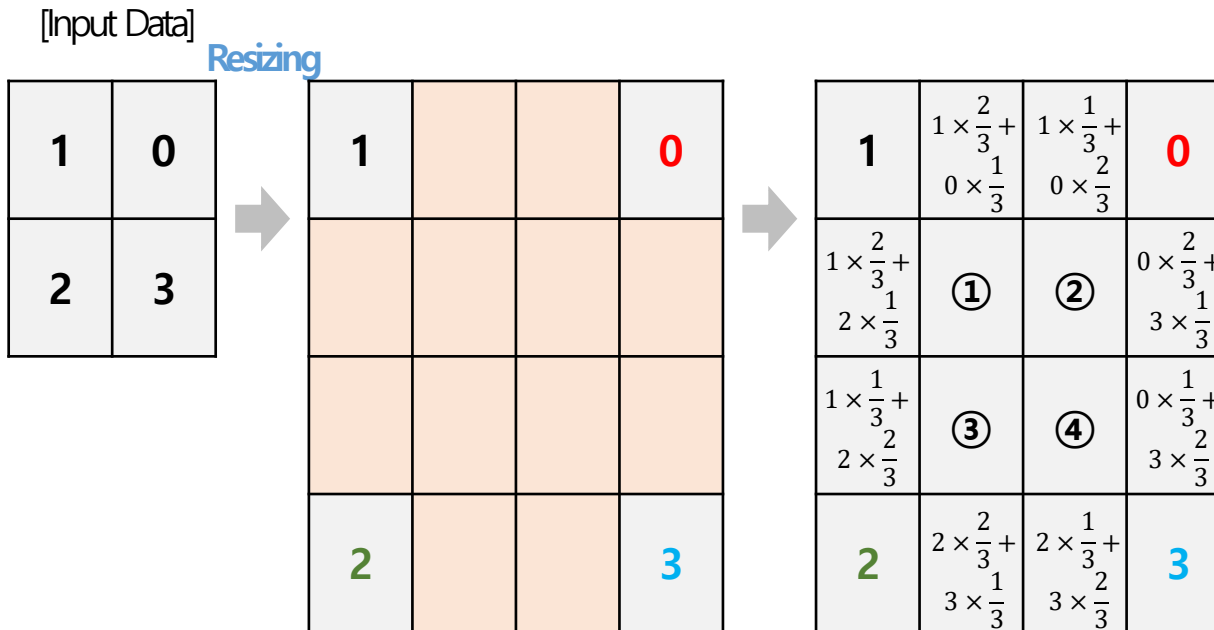
가장 가까운  
픽셀 값으로 채움

1	1	0	0
1	1	0	0
2	2	3	3
2	2	3	3

# Image Super-Resolution

## Bilinear Interpolation

- ❖ 인접한 4개 픽셀에 대한 거리를 가중치로 두어 빈 영역을 채우는 방법
- ❖ 빈 영역에 대해 거리 가중치를 둔 평균값을 계산하는 것과 동일



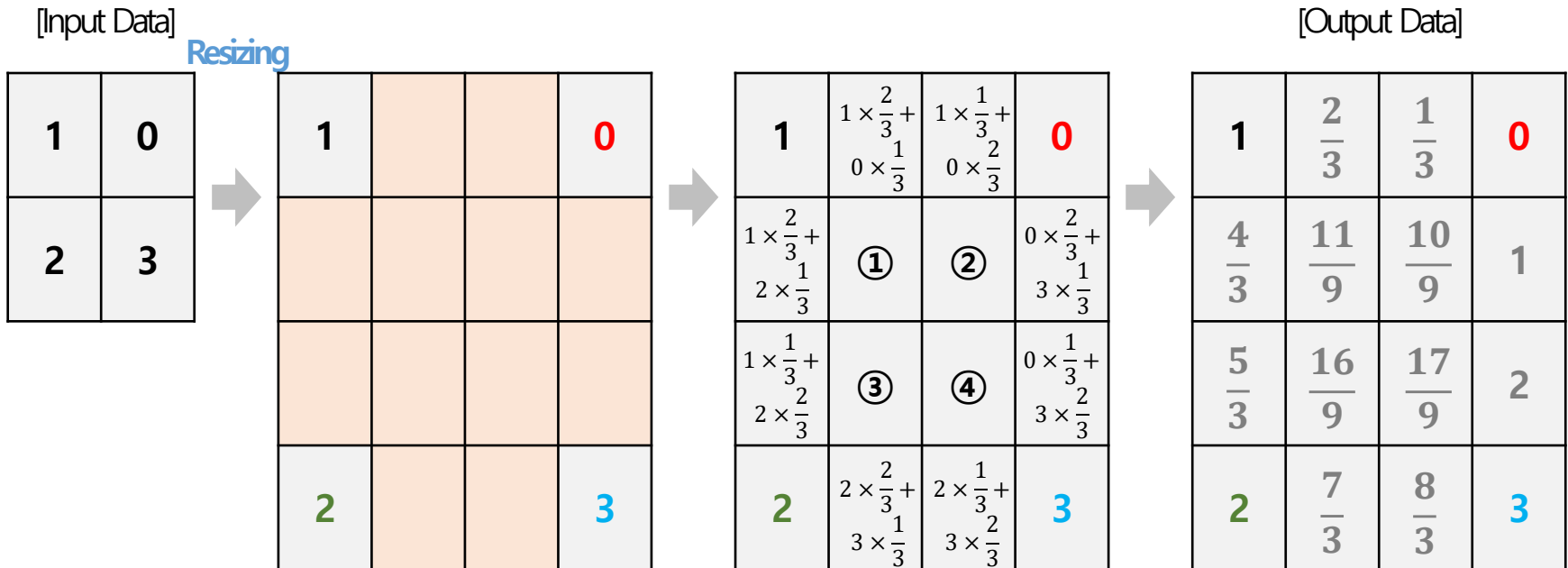
$$\textcircled{1}: 1 \times \frac{4}{9} + 0 \times \frac{2}{9} + 2 \times \frac{2}{9} + 3 \times \frac{1}{9} \quad \textcircled{2}: 1 \times \frac{2}{9} + 0 \times \frac{4}{9} + 2 \times \frac{1}{9} + 3 \times \frac{2}{9}$$

$$\textcircled{3}: 1 \times \frac{2}{9} + 0 \times \frac{1}{9} + 2 \times \frac{4}{9} + 3 \times \frac{2}{9} \quad \textcircled{4}: 1 \times \frac{1}{9} + 0 \times \frac{2}{9} + 2 \times \frac{2}{9} + 3 \times \frac{4}{9}$$

# Image Super-Resolution

## Bilinear Interpolation

- ❖ 인접한 4개 픽셀에 대한 거리를 가중치로 두어 빈 영역을 채우는 방법
- ❖ 빈 영역에 대해 거리 가중치를 둔 평균값을 계산하는 것과 동일



$$\textcircled{1}: 1 \times \frac{4}{9} + 0 \times \frac{2}{9} + 2 \times \frac{2}{9} + 3 \times \frac{1}{9} \quad \textcircled{2}: 1 \times \frac{2}{9} + 0 \times \frac{4}{9} + 2 \times \frac{1}{9} + 3 \times \frac{2}{9}$$

$$\textcircled{3}: 1 \times \frac{2}{9} + 0 \times \frac{1}{9} + 2 \times \frac{4}{9} + 3 \times \frac{2}{9} \quad \textcircled{4}: 1 \times \frac{1}{9} + 0 \times \frac{2}{9} + 2 \times \frac{2}{9} + 3 \times \frac{4}{9}$$

# Image Super-Resolution

## Bicubic Interpolation

- ❖ 인접한 16개 픽셀의 거리에 따른 가중치를 곱해서 계산함
- ❖ 먼저 세로줄 픽셀 값으로 붉은색 영역의 1차원 cubic interpolation을 진행

[Input Data]

1	2	3	4
5	6	7	8
9	10	11	12
13	14	15	16

$$f(x) = \begin{cases} (a + 2)|x|^3 - (a + 3)|x|^2 + 1, & 0 \leq |x| < 1 \\ a|x|^3 - 5a|x|^2 + 8a|x| - 4a, & 1 \leq |x| < 2 \\ 0, & 2 \leq |x| \end{cases}$$

Where,  $a = -0.5, -0.75$  or  $-1.0$

$|x| =$  거리,  $f(x) =$  거리에 따른 가중치

[Output Data]

1			2			3			4
5			6			7			8
9			10			11			12
13			14			15			16

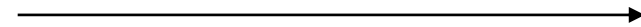
# Image Super-Resolution

## Bicubic Interpolation

- ❖ 인접한 16개 픽셀의 거리에 따른 가중치를 곱해서 계산함
- ❖ 먼저 세로줄 붉은색 영역의 픽셀 값을 1차원 cubic interpolation을 계산
- ❖ 이후 노란색 영역 픽셀 값을 1차원 cubic interpolation으로 계산

[Input Data]

1	2	3	4
5	6	7	8
9	10	11	12
13	14	15	16



$$f(x) = \begin{cases} (a + 2)|x|^3 - (a + 3)|x|^2 + 1, & 0 \leq |x| < 1 \\ a|x|^3 - 5a|x|^2 + 8a|x| - 4a, & 1 \leq |x| < 2 \\ 0, & 2 \leq |x| \end{cases}$$

Where,  $a = -0.5, -0.75$  or  $-1.0$

$|x| =$  거리,  $f(x) =$  거리에 따른 가중치

[Output Data]

1			2			3			4
5			6			7			8
9			10			11			12
13			14			15			16

-인접한 데이터를 가장 많이 활용하는 bicubic interpolation 방법이 가장 성능이 좋고 자주 활용됨-

# Image Super-Resolution

## Interpolation 이미지 비교

- ❖ 원본 이미지의 특정 부분을 확대하면서 다양한 interpolation 방법을 적용
- ❖ Bicubic interpolation 방법일 때 가장 선명하게 확대되었음을 확인



<원본 이미지>



<Nearest Neighbor>



<Bilinear>



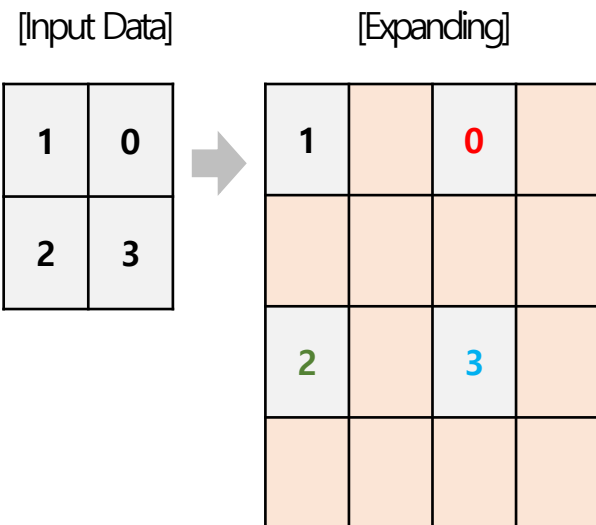
<Bicubic>



# Image Super-Resolution

## Transpose Convolution

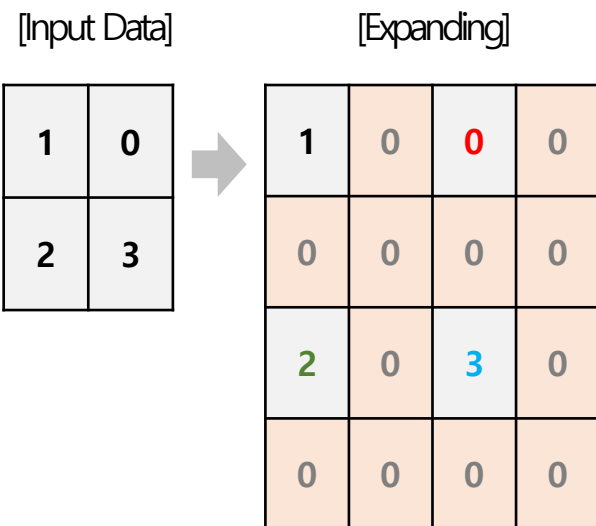
- ❖ Transpose Convolution은 이미지 크기를 확대하기 위해 사용하는 Convolution 기법임
- ❖ 저해상도 이미지를 먼저 고해상도 이미지 크기에 맞춰서 확장함



# Image Super-Resolution

## Transpose Convolution

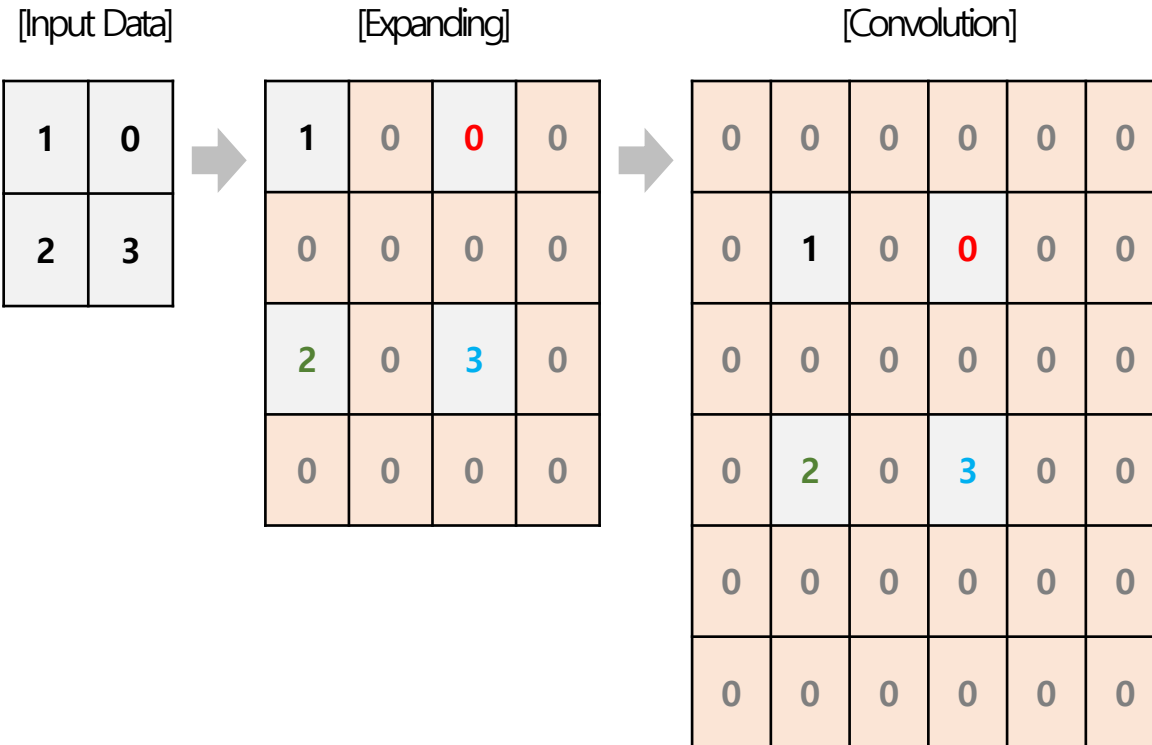
- ❖ 저해상도 이미지를 먼저 고해상도 이미지 크기에 맞춰서 확장함
- ❖ 확장한 영역에는 모두 0 값으로 채움



# Image Super-Resolution

## Transpose Convolution

- ❖ Transpose Convolution 연산 이후 원하는 Output Data 크기로 만들기 위해 필요한 크기만큼 zero padding을 진행

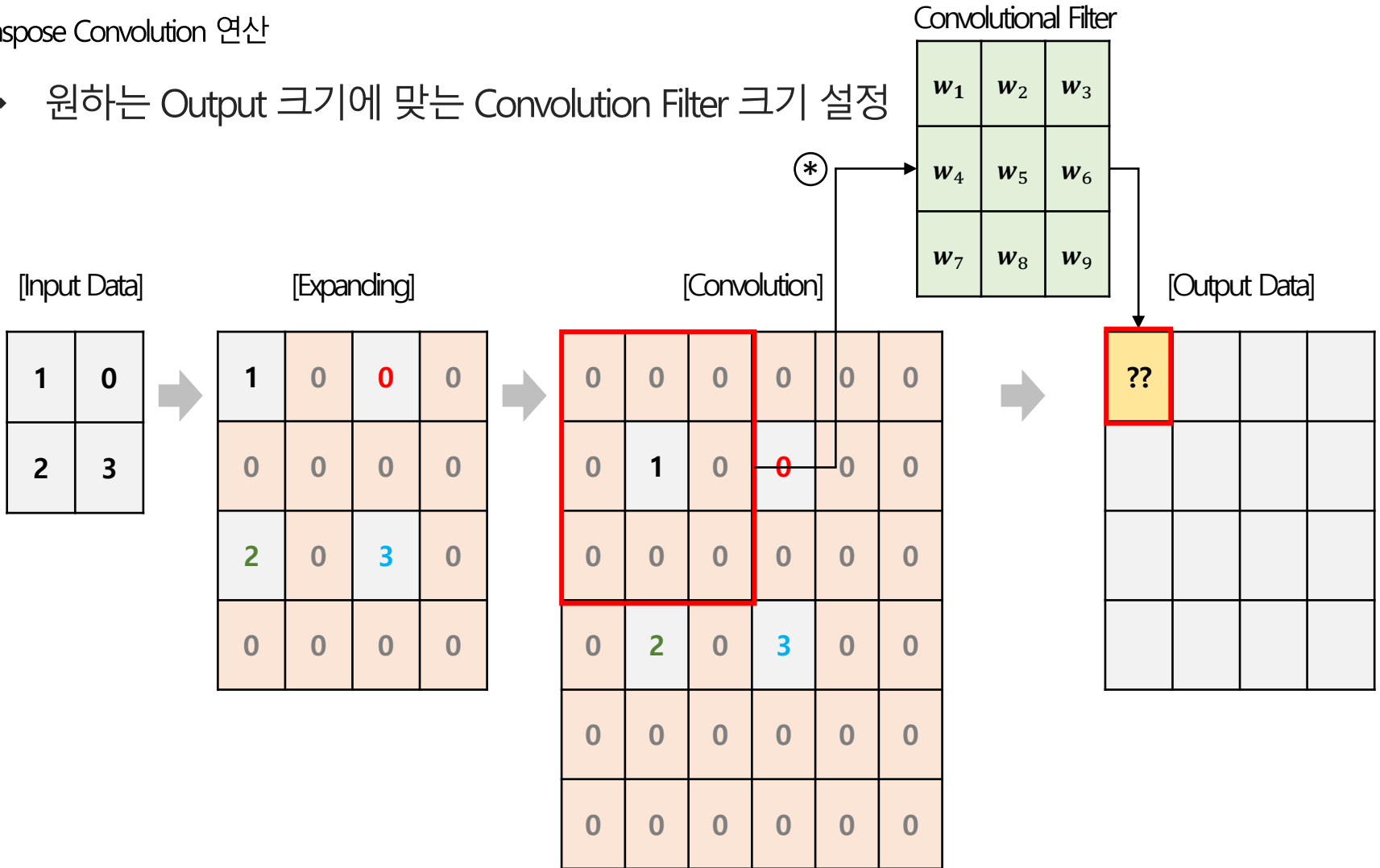


\*\* Output Data 크기에 맞춰 padding

# Image Super-Resolution

Transpose Convolution 연산

- ❖ 원하는 Output 크기에 맞는 Convolution Filter 크기 설정

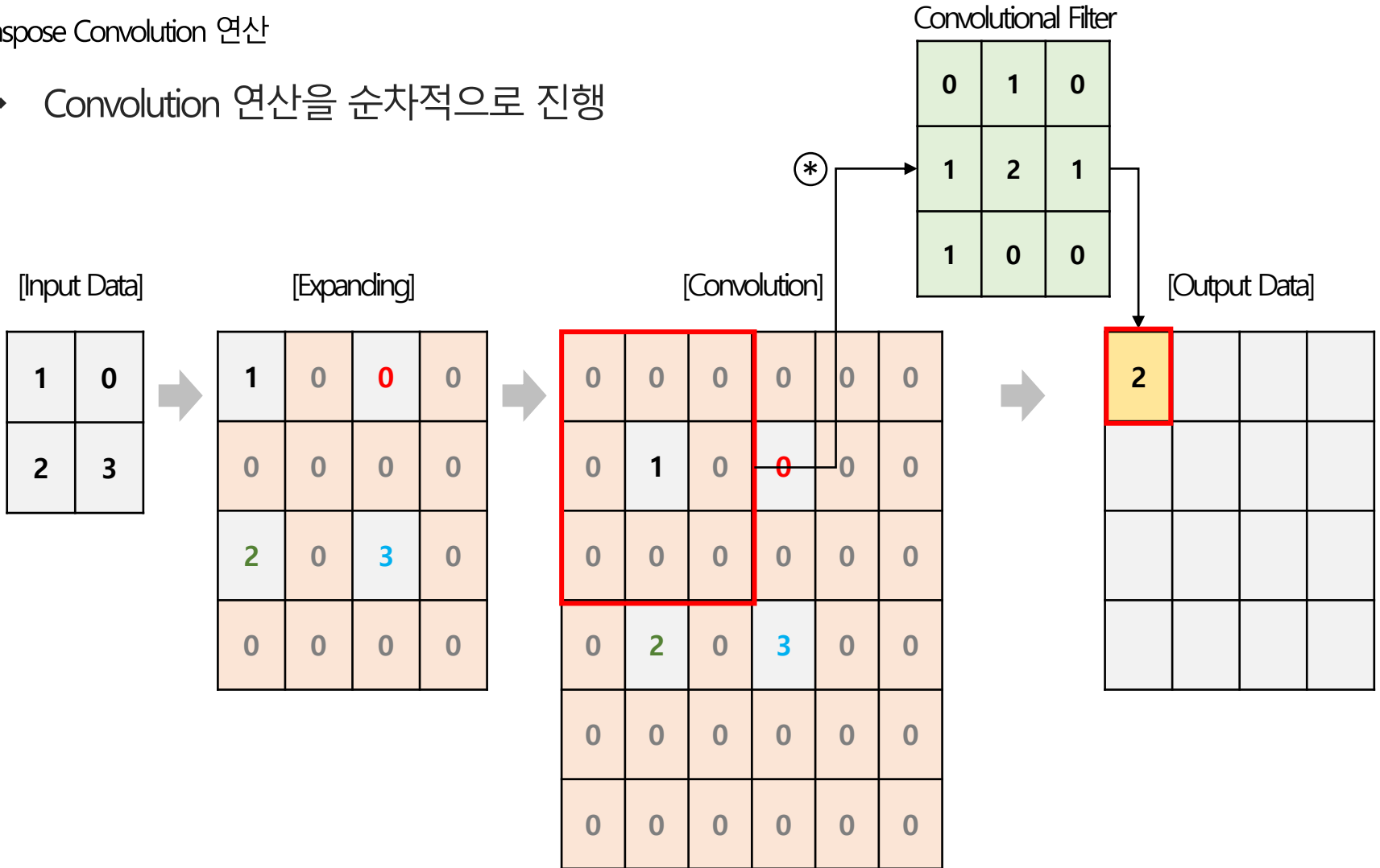


\*\* Output Data 크기에 맞춰 padding

# Image Super-Resolution

Transpose Convolution 연산

- ❖ Convolution 연산을 순차적으로 진행

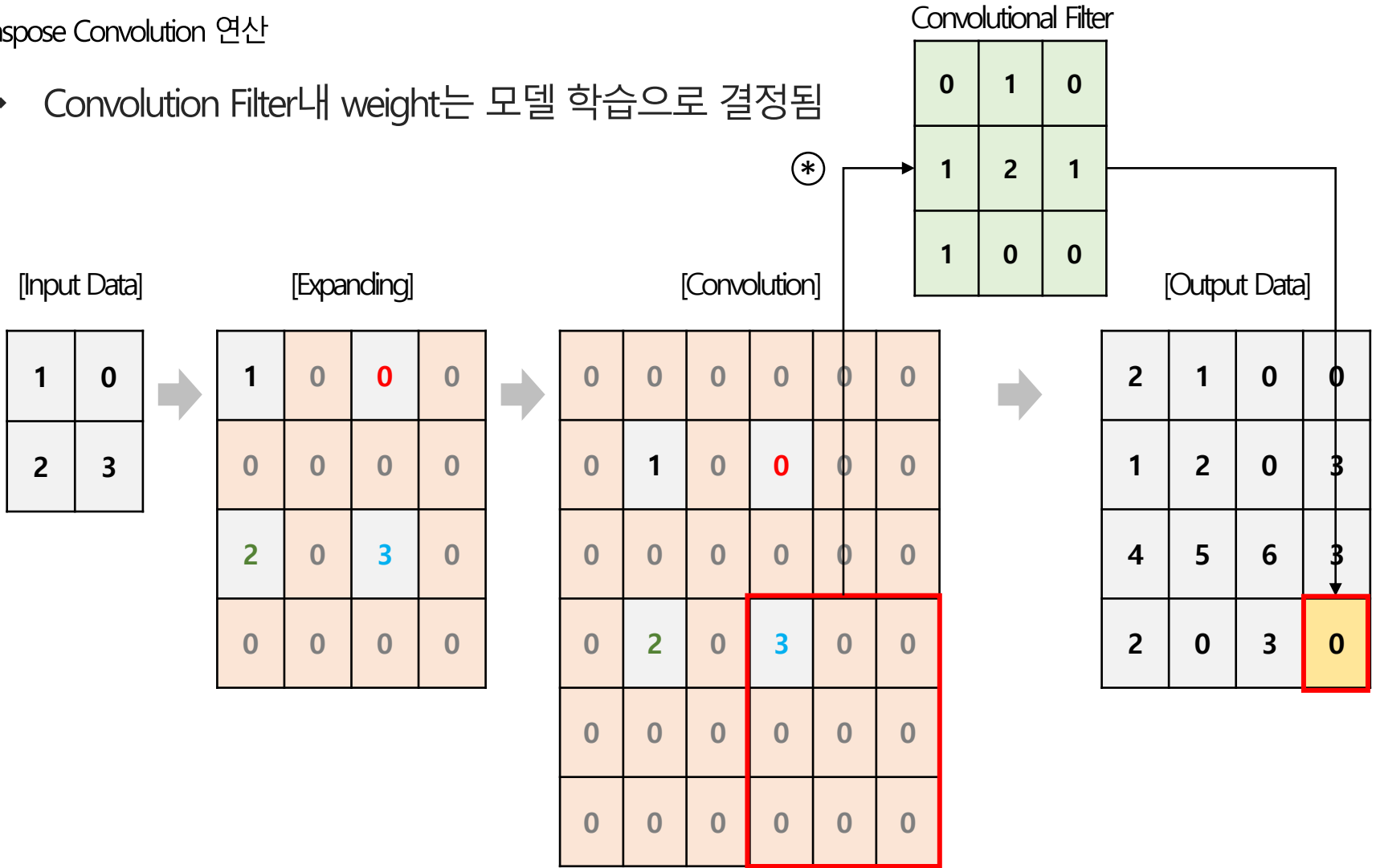


\*\* Output Data 크기에 맞춰 padding

# Image Super-Resolution

Transpose Convolution 연산

❖ Convolution Filter내 weight는 모델 학습으로 결정됨

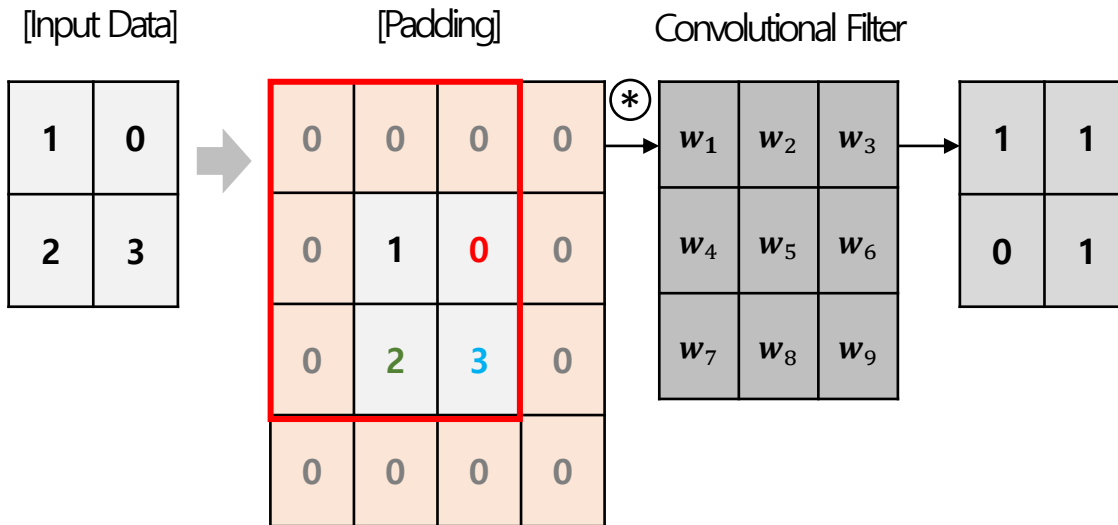


\*\* Output Data 크기에 맞춰 padding

# Image Super-Resolution

## Sub-pixel convolutional layer

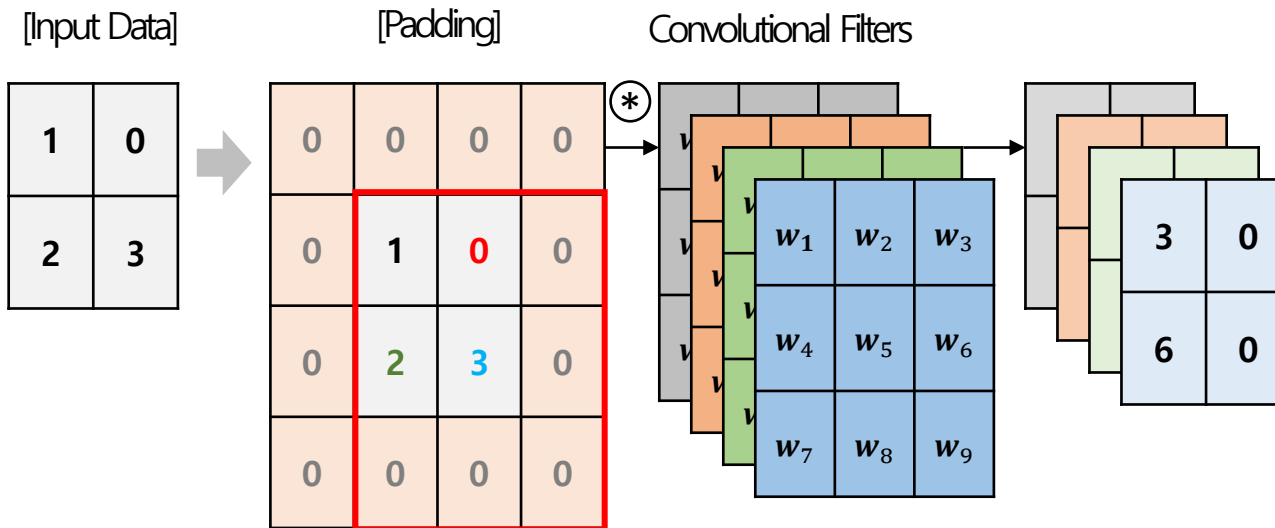
- ❖ Sub-pixel convolutional layer는 convolution 연산으로 여러 개의 feature map을 추출하고 같은 픽셀 위치에 있는 값을 모아 upsampling을 진행
- ❖ 여러 개 convolution 연산을 조합해 성능 향상을 기대할 수 있고, 효율적인 연산도 가능



# Image Super-Resolution

## Sub-pixel convolution 연산

- ❖ Sub-pixel convolutional layer는 convolution 연산으로 여러 개의 feature map을 추출하고 같은 픽셀 위치에 있는 값을 모아 upsampling을 진행

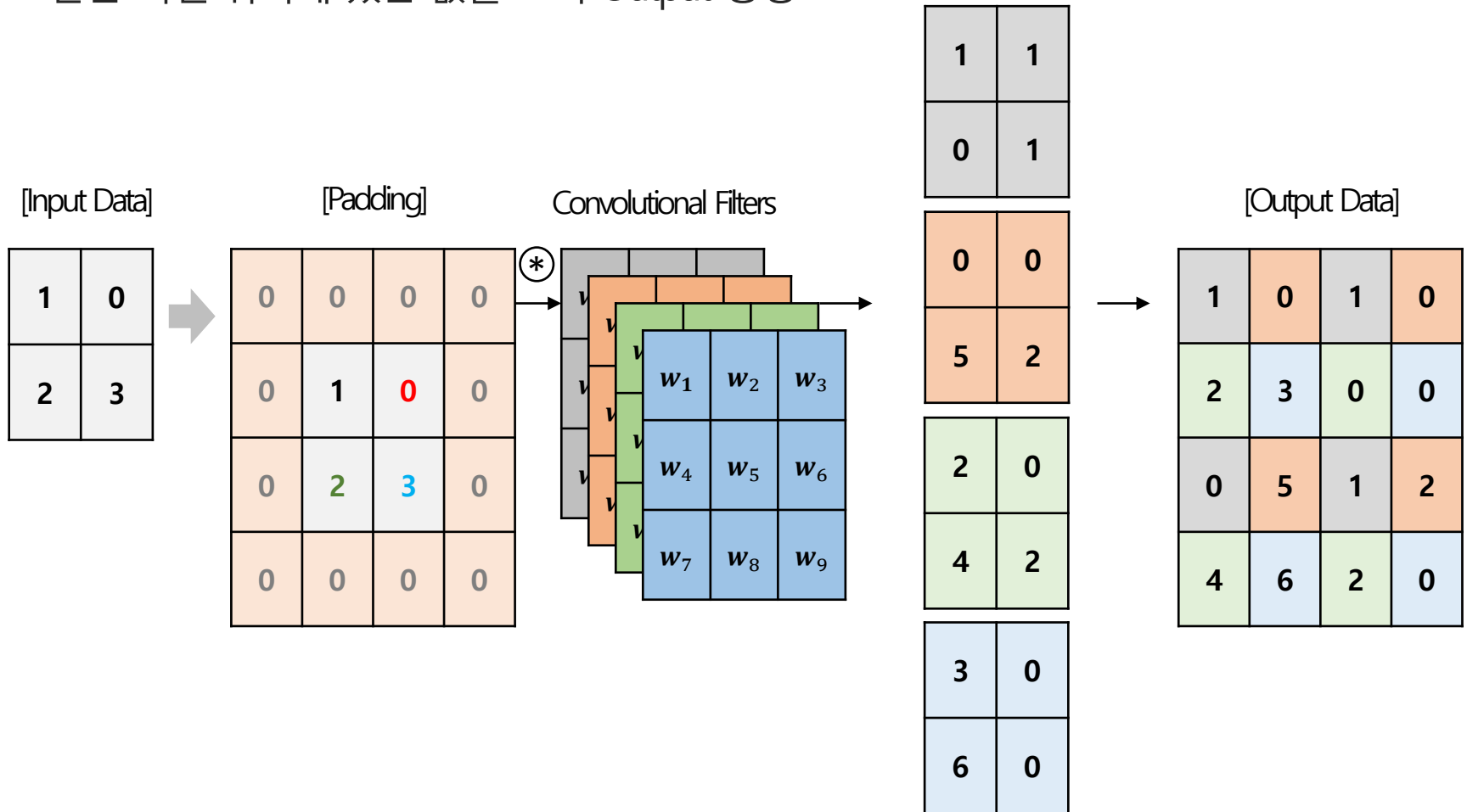




# Image Super-Resolution

## Sub-pixel convolution 연산

- ❖ 같은 픽셀 위치에 있는 값을 모아 Output 생성



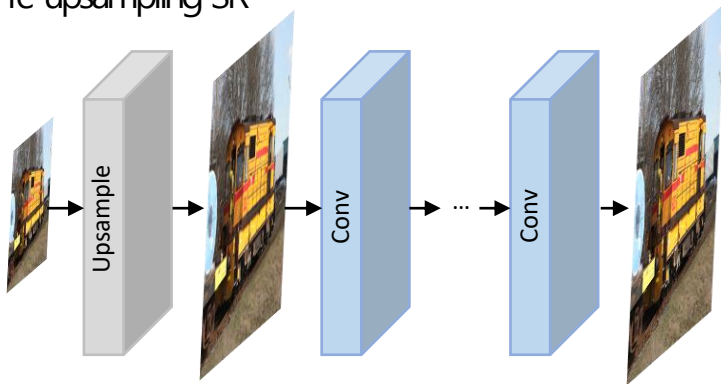
-Transpose convolution, Sub-pixel convolution 모두 학습을 통해 최적 weight를 찾는 방법임-

# Image Super-Resolution

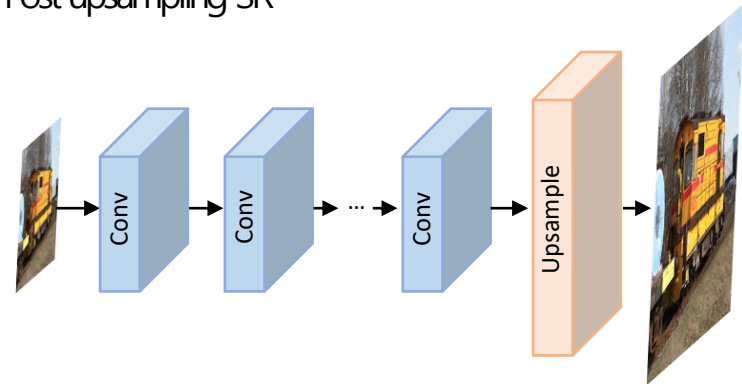
## Image Super-Resolution 프레임워크

- ❖ 모델 내에 Upsampling을 하는 위치에 따라 다양한 Framework가 존재
- ❖ 연구에 따라 Upsampling 방식을 다양하게 적용할 수 있음

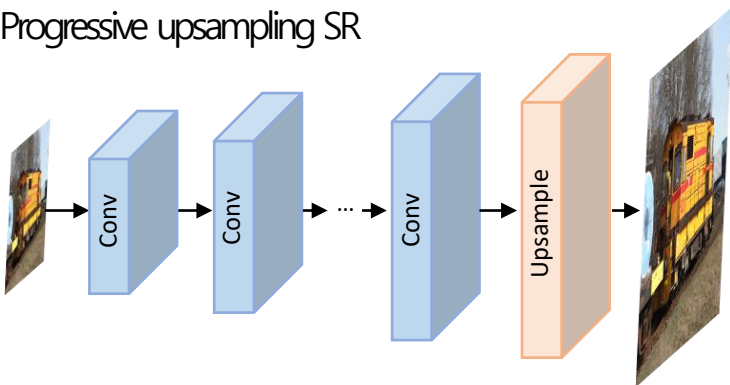
① Pre-upsampling SR



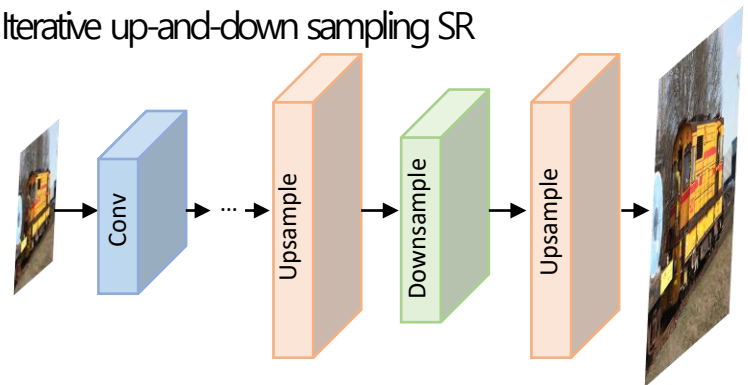
② Post-upsampling SR



③ Progressive upsampling SR



④ Iterative up-and-down sampling SR



## <Deep Learning Frameworks for Super-Resolution>

출처: Wang, Z., Chen, J., & Hoi, S. C. (2020). Deep learning for image super-resolution: A survey. *IEEE transactions on pattern analysis and machine intelligence*.

# Image Super-Resolution

---

## Image Super-Resolution 특징

1. 저해상도 이미지와 고해상도 이미지 크기가 다른 문제를 어떻게 다룰 것인가?  
→ Interpolation, Transpose convolution, Sub-pixel convolution 등으로 Upsampling 진행
2. Image Super-Resolution 문제에 어떻게 딥러닝 모델을 적용할 것인가?
3. 고해상도 이미지 변환이 잘 되었다는 것을 어떻게 평가할 것인가?

# Image Super-Resolution

---

## Image Super-Resolution 특징

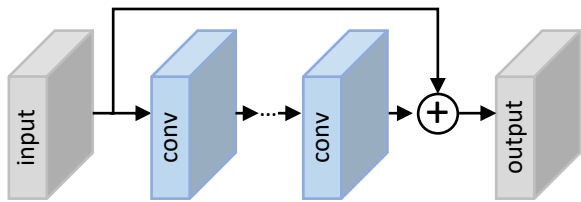
1. 저해상도 이미지와 고해상도 이미지 크기가 다른 문제를 어떻게 다룰 것인가?  
→ Interpolation, Transpose Convolution, Sub-pixel 방법 등으로 Upsampling 진행
2. Image Super-Resolution 문제에 어떻게 딥러닝 모델을 적용할 것인가?  
→ 일반적인 분류, 회귀 모델에서 사용하는 모든 방법론이 적용 가능함
3. 고해상도 이미지 변환이 잘 되었다는 것을 어떻게 평가할 것인가?

# Image Super-Resolution

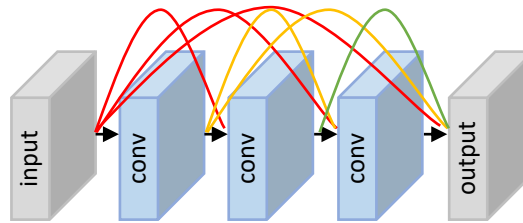
## Image Super-Resolution 네트워크

- ❖ ResNet, DenseNet 등 일반적인 분류 문제에서 사용한 딥러닝 모델 구조 및 특징 적용
- ❖ Input 이미지의 특징을 잘 추출하기 위한 Network 구성도 있음

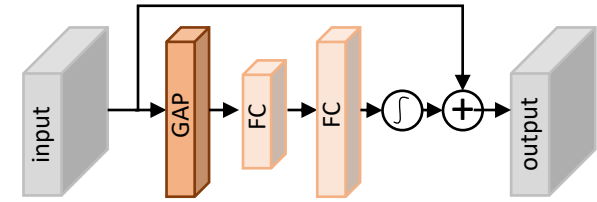
① Residual Learning



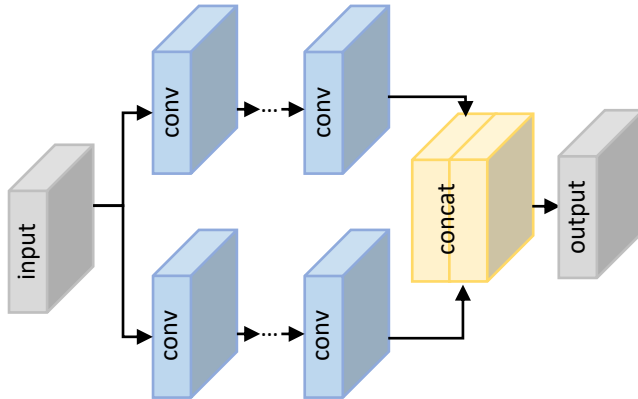
② Dense connection



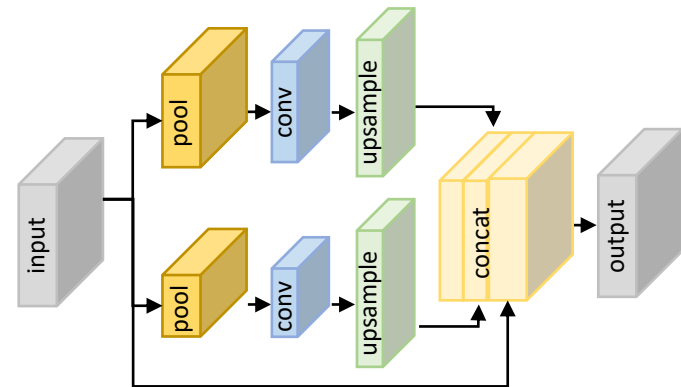
③ Channel attention



④ Local multi-path learning



⑤ Pyramid pooling



<Network design strategies for image super-resolution>

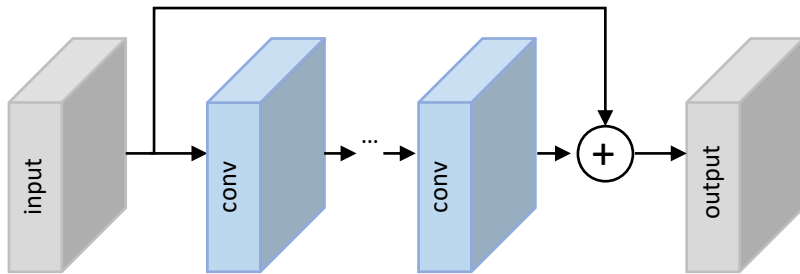
출처: Wang, Z., Chen, J., & Hoi, S. C. (2020). Deep learning for image super-resolution: A survey. *IEEE transactions on pattern analysis and machine intelligence*.

# Image Super-Resolution

일반적인 분류 문제에서 사용하는 딥러닝 모델 구조 및 특징

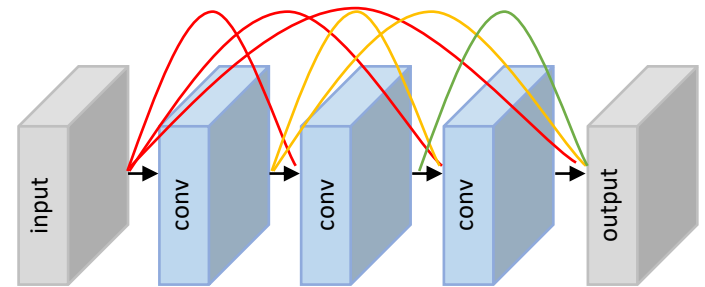
- ❖ Image Super-Resolution 모델 성능을 향상시키기 위해 일반적인 분류 문제에서 사용하는 딥러닝 모델 구조를 다양하게 적용한 연구들이 많이 진행되고 있음

① Residual Learning



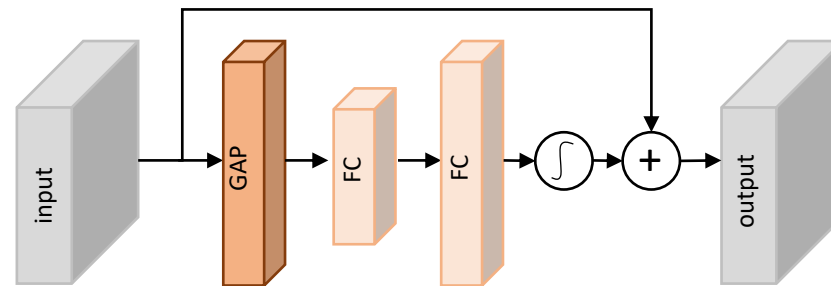
→ Input 정보를 feature 정보에 더하면서 학습 진행

② Dense connection



→ 이전 feature 정보를 모두 concatenate 하며 학습 진행

③ Channel attention

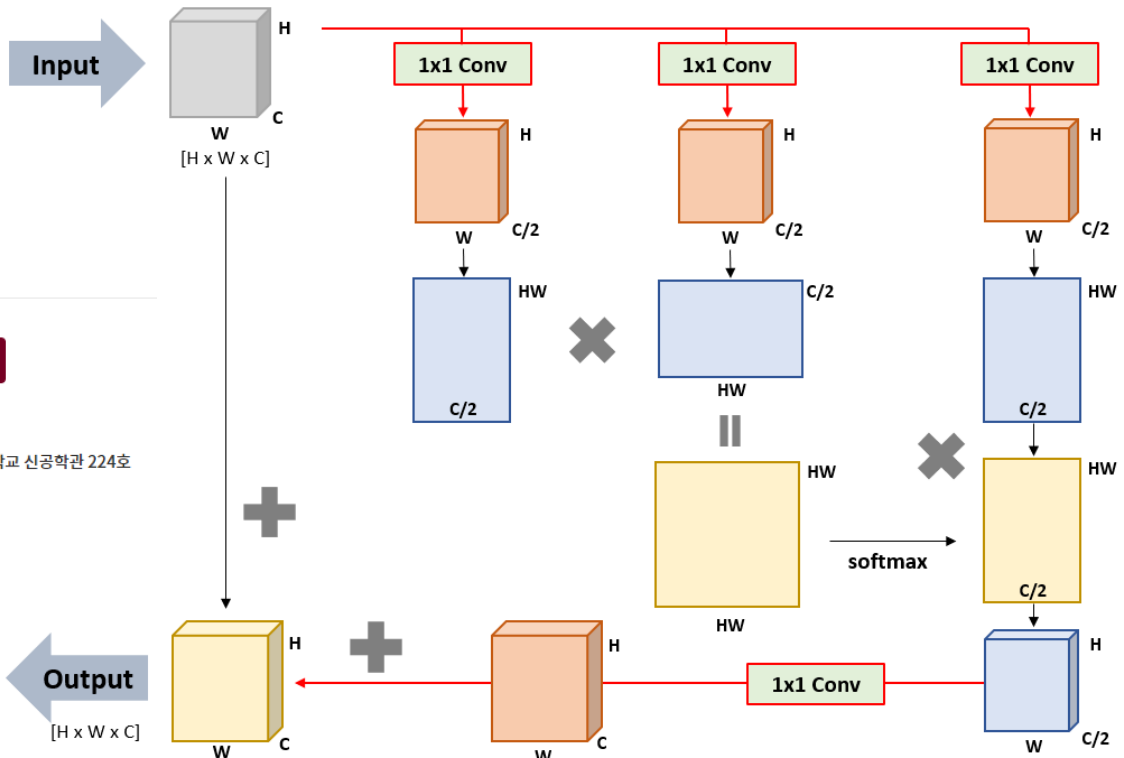


→ channel 단위의 관계 정보에 집중하며 학습을 진행

# Image Super-Resolution

## Non-local neural networks

- ❖ 1x1 Convolutional Filter를 활용해 Input의 특징을 다양하게 추출
- ❖ Feature 내 픽셀간 관계 정보를 고려해 학습하는 모듈임



## Non-local Neural Networks

2020년 7월 17일 오후 1:42 / 조회수: 1162

### REFERENCES

20200724\_DMQA세미나\_백인성\_최종본.pdf

### INFORMATION

2020년 7월 24일 오후 1시 ~ 고려대학교 신공학관 224호

발표자:  백인성

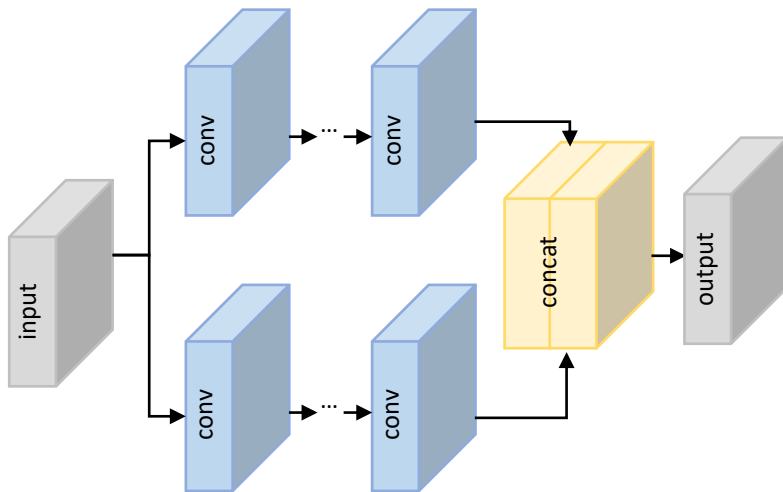
출처: Wang, X., Girshick, R., Gupta, A., & He, K. (2018). Non-local neural networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 7794-7803). <http://dmqa.korea.ac.kr/activity/seminar/290>

# Image Super-Resolution

## Multi-path learning & Pyramid pooling

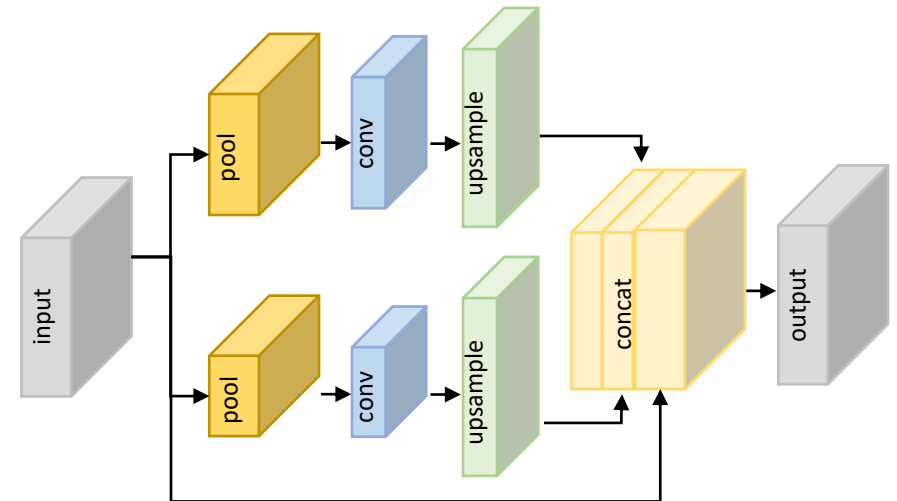
- ❖ Local multi-path learning, Pyramid pooling 모두 input 정보에서 좋은 특징을 다양하게 추출할 수 있는 방안을 제시하고 있음

④ Local multi-path learning



→ Input 정보에서 여러 개 feature를 도출해 결합해 학습 진행

⑤ Pyramid pooling



→ global한 특징과 local한 특징을 잘 추출해 학습 진행



# Image Super-Resolution

## Loss function in Image Super-Resolution

- ❖ 많은 연구에서 L1 loss와 L2 loss를 사용하고 있음

명칭	수식	특징
L1 loss	$L_{pixel\_l1}(\hat{I}, I) = \frac{1}{hwc} \sum_{i,j,k}  \hat{I}_{i,j,k} - I_{i,j,k} $	픽셀별 Mean Absolute Error(MAE) 값과 같음
L2 loss	$L_{pixel\_l2}(\hat{I}, I) = \frac{1}{hwc} \sum_{i,j,k} (\hat{I}_{i,j,k} - I_{i,j,k})^2$	픽셀별 Mean Square Error(MSE) 값과 같음
Charbonnier loss	$L_{pixel\_cha}(\hat{I}, I) = \frac{1}{hwc} \sum_{i,j,k} \sqrt{(\hat{I}_{i,j,k} - I_{i,j,k})^2 + \epsilon^2}$	Penalty term( $\epsilon$ )을 넣어 Outlier에 robust함
Content loss	$L_{content}(\hat{I}, I; \phi, l) = \frac{1}{h_l w_l c_l wc} \sqrt{\sum_{i,j,k} (\phi_{i,j,k}^{(l)}(\hat{I}) - \phi_{i,j,k}^{(l)}(I))^2}$	Image의 자세한 특징을 잘 잡아내기 위해 제시된 Loss임

*I*: target HR image  
*Ŷ*: generated HR image  
*h*: height  
*w*: width

*c*: number of channels  
*i, j, k*: pixel index  
 $\epsilon$ : constant  
 $\phi^l(I)$ : high-level representations on *l*-th layer

# Image Super-Resolution

---

## Image Super-Resolution 특징

1. 저해상도 이미지와 고해상도 이미지 크기가 다른 문제를 어떻게 다룰 것인가?  
→ Interpolation, Transpose Convolution, Sub-pixel 방법 등으로 Upsampling 진행
2. Image Super-Resolution 문제에 어떻게 딥러닝 모델을 적용할 것인가?  
→ 일반적인 딥러닝 모델에서 사용하는 알고리즘 모두 사용 가능
3. 고해상도 이미지 변환이 잘 되었다는 것을 어떻게 평가할 것인가?

# Image Super-Resolution

## PSNR

- ❖ Peak signal to noise ratio(PSNR)은 최대 신호 대 잡음비를 의미
- ❖ 압축된 영상의 화질에 대한 손실 정보를 평가

$$PSNR = 10 \cdot \log_{10} \left( \frac{MAX_I}{\frac{1}{N} \sum_{i=1}^N (I(i) - \hat{I}(i))^2} \right) = 10 \times \log_{10} \left( \frac{MAX_I^2}{MSE} \right)$$

$MAX_I$  = 일반적인 이미지에서 255 값을 의미

$$MSE = \frac{1}{N} \sum_{i=1}^N (I(i) - \hat{I}(i))^2$$

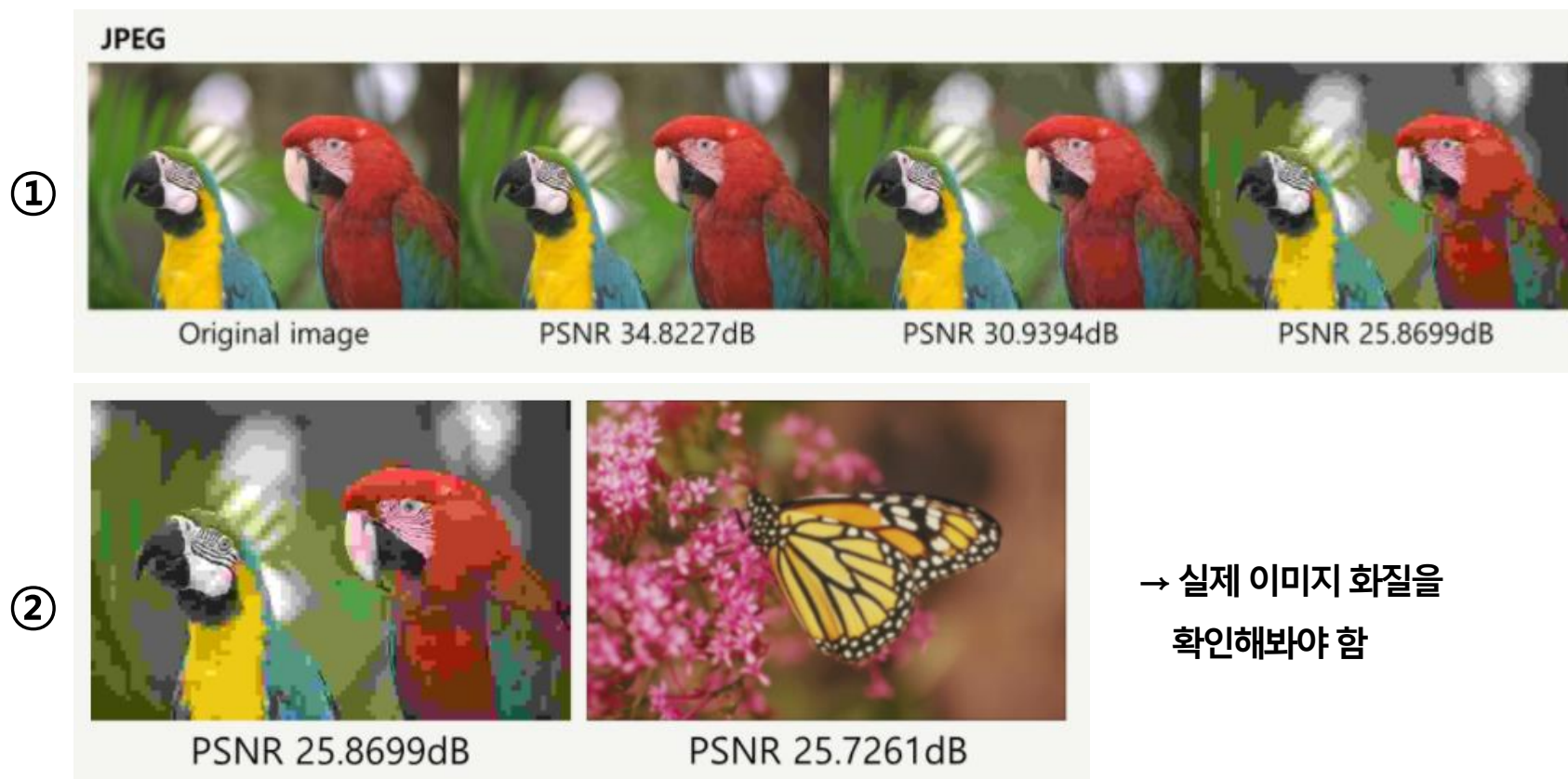
예측 이미지 픽셀값

원본 이미지 픽셀값

# Image Super-Resolution

## PSNR

- ❖ PSNR은 원본 이미지 대비 손실된 품질의 정도를 알 수 있음
- ❖ 하지만 원본 이미지에 따라 PSNR 값이 유사하더라도 화질의 차이 있을 수 있음



출처: <https://bskyvision.com/392>

# Image Super-Resolution

## SSIM

- ❖ Structural similarity index(SSIM)은 인간의 시각적 화질 차이 및 유사도 평가 위해 고안
- ❖ 이미지의 구조 정보의 왜곡 정도가 사람의 시각에 큰 영향을 미친다는 것을 반영

[휘도]

$$C_l(I, \hat{I}) = \frac{2\mu_I\mu_{\hat{I}} + C_1}{\mu_I^2 + \mu_{\hat{I}}^2 + C_1}$$

[대비]

$$C_c(I, \hat{I}) = \frac{2\sigma_I\sigma_{\hat{I}} + C_1}{\sigma_I^2 + \sigma_{\hat{I}}^2 + C_1}$$

[구조]

$$C_s(I, \hat{I}) = \frac{\sigma_{I\hat{I}} + C_3}{\sigma_I\sigma_{\hat{I}} + C_1}$$

$$C_1=(k_1L)^2, \quad C_2=(k_2L)^2, \quad k_1, k_2 \ll 1, \quad C_1, C_2, C_3: \text{constant}, \quad \sigma_I\sigma_{\hat{I}}: \text{covariance between } I \text{ and } \hat{I}$$

$I$ : target HR image,       $\hat{I}$ : generated HR image

$$SSIM(I, \hat{I}) = [C_l(I, \hat{I})]^\alpha [C_c(I, \hat{I})]^\beta [C_s(I, \hat{I})]^\gamma$$

$\alpha, \beta, \gamma = \text{hyperparameters}$

# Image Super-Resolution

---

## Image Super-Resolution 특징

1. 저해상도 이미지와 고해상도 이미지 크기가 다른 문제를 어떻게 다룰 것인가?  
→ Interpolation, Transpose Convolution, Sub-pixel 방법 등이 존재
2. Image Super-Resolution 문제에 어떻게 딥러닝 모델을 적용할 것인가?  
→ 일반적인 딥러닝 모델에서 사용하는 알고리즘 모두 사용 가능
3. 고해상도 이미지 변환이 잘 되었다는 것을 어떻게 평가할 것인가?  
→ PSNR, SSIM을 중심으로 고해상도 이미지 변환이 잘 되었음을 평가

# Image Super-Resolution

## Image Super-Resolution 모델 정리

- ❖ 모델별 upsampling 방식, 네트워크 구조 특징, loss function, keywords가 정리되어 있음

TABLE 2

Super-resolution methodology employed by some representative models. The "Fw.", "Up.", "Rec.", "Res.", "Dense.", "Att." represent SR frameworks, upsampling methods, recursive learning, residual learning, dense connections, attention mechanism, respectively.

Method	Publication	Fw.	Up.	Rec.	Res.	Dense	Att.	$\mathcal{L}_{L1}$	$\mathcal{L}_{L2}$	Keywords
SRCNN [22]	2014, ECCV	Pre.	Bicubic						✓	
DRCN [82]	2016, CVPR	Pre.	Bicubic	✓	✓				✓	Recursive layers
FSRCNN [43]	2016, ECCV	Post.	Deconv						✓	Lightweight design
ESPCN [156]	2017, CVPR	Pre.	Sub-Pixel						✓	Sub-pixel
LapSRN [27]	2017, CVPR	Pro.	Bicubic		✓			✓		$\mathcal{L}_{\text{pixel\_Cha}}$
DRRN [56]	2017, CVPR	Pre.	Bicubic	✓	✓				✓	Recursive blocks
SRResNet [25]	2017, CVPR	Post.	Sub-Pixel		✓				✓	$\mathcal{L}_{\text{Con.}}, \mathcal{L}_{TV}$
SRGAN [25]	2017, CVPR	Post.	Sub-Pixel		✓					$\mathcal{L}_{\text{Con.}}, \mathcal{L}_{\text{GAN}}$
EDSR [31]	2017, CVPRW	Post.	Sub-Pixel		✓			✓		Compact and large-size design
EnhanceNet [8]	2017, ICCV	Pre.	Bicubic		✓					$\mathcal{L}_{\text{Con.}}, \mathcal{L}_{\text{GAN}}, \mathcal{L}_{\text{texture}}$
MemNet [55]	2017, ICCV	Pre.	Bicubic	✓	✓	✓			✓	Memory block
SRDenseNet [79]	2017, ICCV	Post.	Deconv		✓	✓			✓	Dense connections
DBPN [57]	2018, CVPR	Iter.	Deconv		✓	✓			✓	Back-projection
DSRN [85]	2018, CVPR	Pre.	Deconv	✓	✓				✓	Dual state
RDN [93]	2018, CVPR	Post.	Sub-Pixel		✓	✓		✓		Residual dense block
CARN [28]	2018, ECCV	Post.	Sub-Pixel	✓	✓	✓		✓		Cascading
MSRN [99]	2018, ECCV	Post.	Sub-Pixel		✓			✓		Multi-path
RCAN [70]	2018, ECCV	Post.	Sub-Pixel		✓		✓	✓		Channel attention
ESRGAN [103]	2018, ECCVW	Post.	Sub-Pixel		✓	✓		✓		$\mathcal{L}_{\text{Con.}}, \mathcal{L}_{\text{GAN}}$
RNAN [106]	2019, ICLR	Post.	Sub-Pixel		✓		✓	✓		Non-local attention
Meta-RDN [95]	2019, CVPR	Post.	Meta Upscale		✓	✓		✓		Magnification-arbitrary
SAN [105]	2019, CVPR	Post.	Sub-Pixel		✓		✓	✓		Second-order attention
SRFBN [86]	2019, CVPR	Post.	Deconv	✓	✓	✓		✓		Feedback mechanism

출처: Wang, Z., Chen, J., & Hoi, S. C. (2020). Deep learning for image super-resolution: A survey. *IEEE transactions on pattern analysis and machine intelligence*.

---

# 3. Image Super-Resolution using Deep Convolutional Networks (SRCNN)

---



- ❖ 2016년 IEEE transactions on pattern analysis and machine intelligence에 논문 게재
- ❖ 2021년 4월 8일 기준 4400회 인용

## Image Super-Resolution Using Deep Convolutional Networks

Chao Dong, Chen Change Loy, *Member, IEEE*, Kaiming He, *Member, IEEE*, and Xiaou Tang, *Fellow, IEEE*

**Abstract**—We propose a deep learning method for single image super-resolution (SR). Our method directly learns an end-to-end mapping between the low/high-resolution images. The mapping is represented as a deep convolutional neural network (CNN) that takes the low-resolution image as the input and outputs the high-resolution one. We further show that traditional sparse-coding-based SR methods can also be viewed as a deep convolutional network. But unlike traditional methods that handle each component separately, our method jointly optimizes all layers. Our deep CNN has a lightweight structure, yet demonstrates state-of-the-art restoration quality, and achieves fast speed for practical on-line usage. We explore different network structures and parameter settings to achieve trade-offs between performance and speed. Moreover, we extend our network to cope with three color channels simultaneously, and show better overall reconstruction quality.

**Index Terms**—Super-resolution, deep convolutional neural networks, sparse coding

### Image super-resolution using deep convolutional networks

[C Dong, CC Loy, K He, X Tang - IEEE transactions on pattern ..., 2015 - ieeeexplore.ieee.org](#)

We propose a deep learning method for single image super-resolution (SR). Our method directly learns an end-to-end mapping between the low/high-resolution images. The mapping is represented as a deep convolutional neural network (CNN) that takes the low ...

☆ 4400회 인용 관련 학술자료 전체 11개의 버전

# SRCNN

## SRCNN model architecture

- ❖ SRCNN은 Image super-resolution 문제에 최초로 deep learning을 적용한 연구임
- ❖ Convolutional Layer 3개로 모델을 구축함, 마지막 Layer에는 ReLU 함수 사용하지 않음

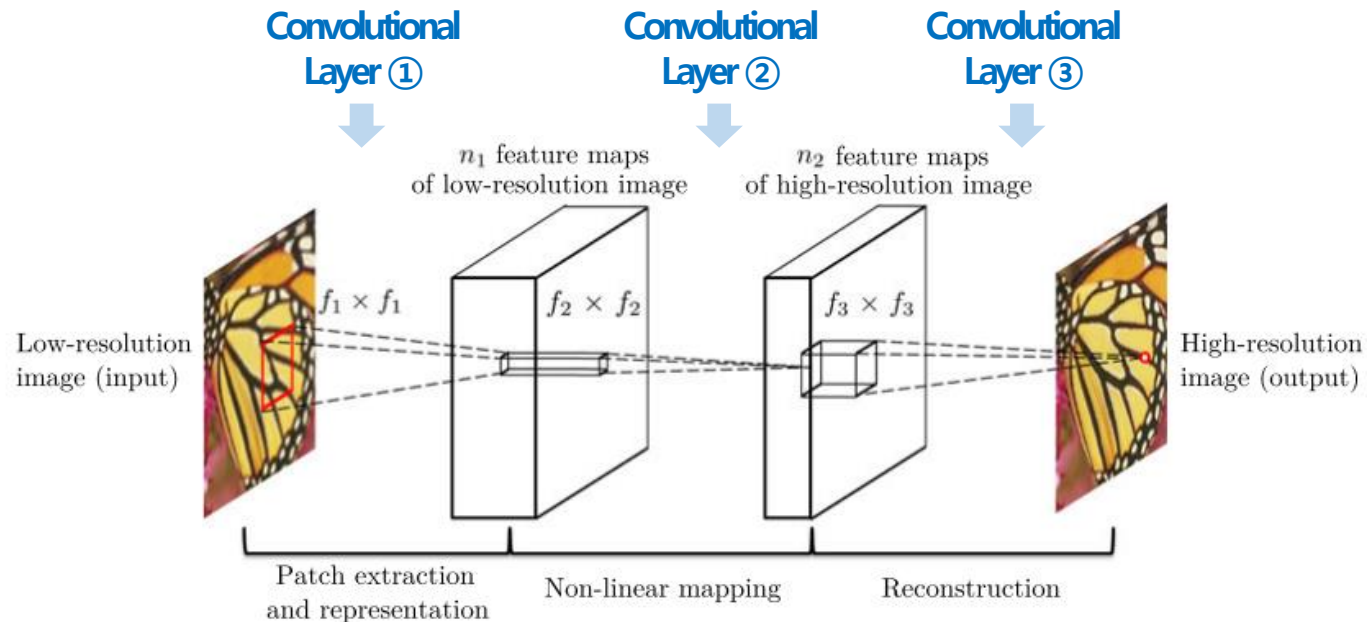


Fig. 2. Given a low-resolution image  $Y$ , the first convolutional layer of the SRCNN extracts a set of feature maps. The second layer maps these feature maps nonlinearly to high-resolution patch representations. The last layer combines the predictions within a spatial neighbourhood to produce the final high-resolution image  $F(Y)$ .

## SRCNN 정량 평가

- ❖ PSNR, SSIM을 포함해서 총 6개 평가지표를 기반으로 성능을 평가
- ❖ PSNR, SSIM을 기준으로 평가했을 때, 제안한 SRCNN이 가장 좋은 성능을 보였음

TABLE 2  
The Average Results of PSNR (dB), SSIM, IFC, NQM, WPSNR (dB) and MSSIM on the Set5 Dataset

Eval. Mat	Scale	Bicubic	SC [48]	NE+LLE [4]	KK [24]	ANR [39]	A+ [39]	SRCNN
PSNR	2	33.66	-	35.77	36.20	35.83	36.54	<b>36.66</b>
	3	30.39	31.42	31.84	32.28	31.92	32.59	<b>32.75</b>
	4	28.42	-	29.61	30.03	29.69	30.28	<b>30.49</b>
SSIM	2	0.9299	-	0.9490	0.9511	0.9499	<b>0.9544</b>	0.9542
	3	0.8682	0.8821	0.8956	0.9033	0.8968	0.9088	<b>0.9090</b>
	4	0.8104	-	0.8402	0.8541	0.8419	0.8603	<b>0.8628</b>
IFC	2	6.10	-	7.84	6.87	8.09	<b>8.48</b>	8.05
	3	3.52	3.16	4.40	4.14	4.52	<b>4.84</b>	4.58
	4	2.35	-	2.94	2.81	3.02	<b>3.26</b>	3.01
NQM	2	36.73	-	42.90	39.49	43.28	<b>44.58</b>	41.13
	3	27.54	27.29	32.77	32.10	33.10	<b>34.48</b>	33.21
	4	21.42	-	25.56	24.99	25.72	<b>26.97</b>	25.96
WPSNR	2	50.06	-	58.45	57.15	58.61	<b>60.06</b>	59.49
	3	41.65	43.64	45.81	46.22	46.02	<b>47.17</b>	47.10
	4	37.21	-	39.85	40.40	40.01	41.03	<b>41.13</b>
MSSSIM	2	0.9915	-	0.9953	0.9953	0.9954	<b>0.9960</b>	0.9959
	3	0.9754	0.9797	0.9841	0.9853	0.9844	<b>0.9867</b>	0.9866
	4	0.9516	-	0.9666	0.9695	0.9672	0.9720	<b>0.9725</b>

# SRCNN

## SRCNN 정성 평가

- ❖ 실제 이미지를 확인했을 때도 가장 선명하게 이미지 복원 되었음을 알 수 있음

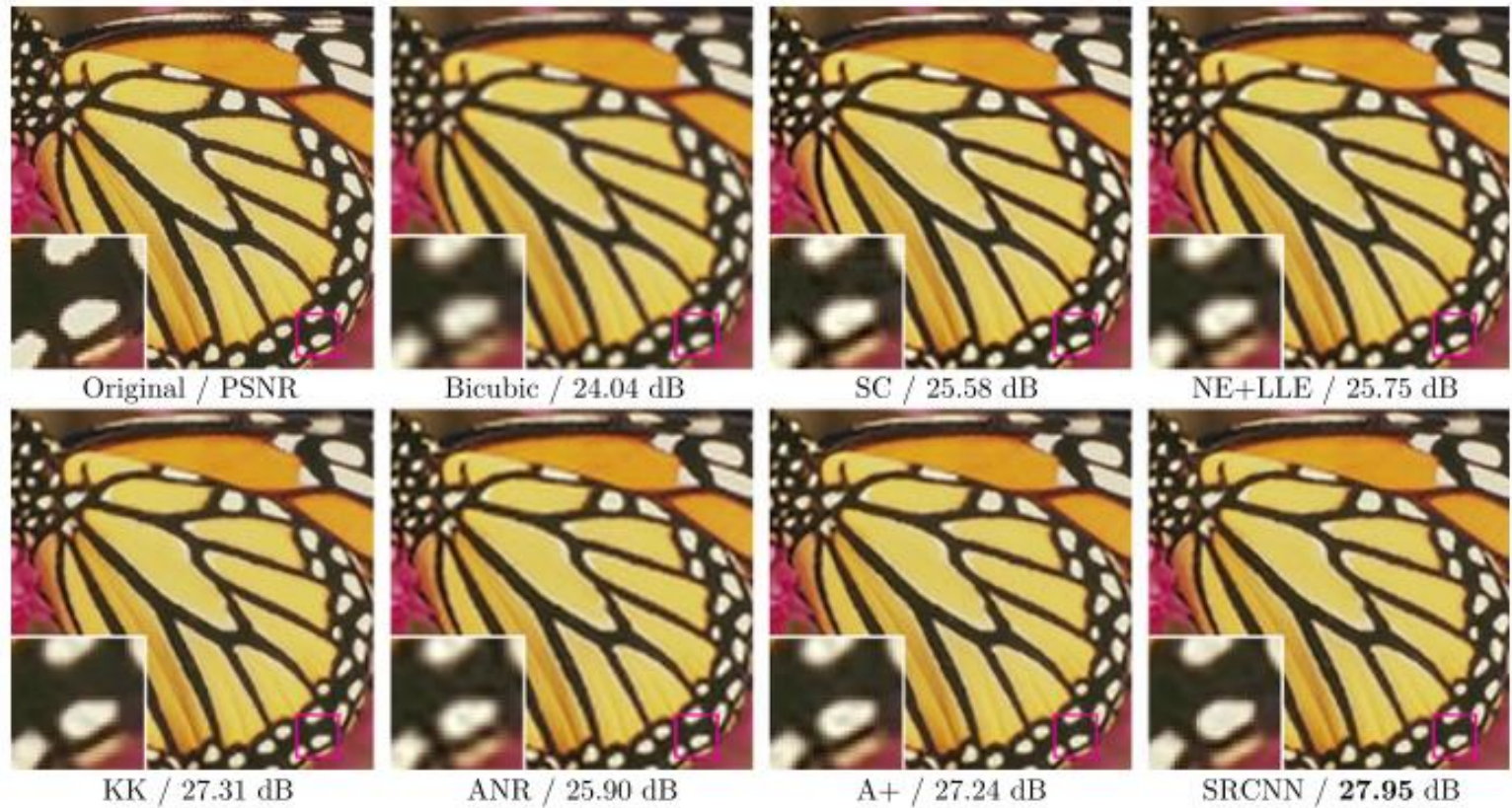


Fig. 14. The “butterfly” image from Set5 with an upscaling factor 3.

출처: Dong, C., Loy, C. C., He, K., & Tang, X. (2015). Image super-resolution using deep convolutional networks. *IEEE transactions on pattern analysis and machine intelligence*, 38(2), 295-307.

---

# 4. Second-order Attention Network for Single Image Super-Resolution (SAN)

---

- ❖ 2019 CVPR(Computer Vision and Pattern Recognition)에서 발표됨
- ❖ 2021년 4월 8일 기준 275회 인용

### Second-order Attention Network for Single Image Super-Resolution

Tao Dai<sup>1,2,\*†</sup>, Jianrui Cai<sup>3,\*</sup>, Yongbing Zhang<sup>1</sup>, Shu-Tao Xia<sup>1,2</sup>, Lei Zhang<sup>3,4,§</sup>

<sup>1</sup>Graduate School at Shenzhen, Tsinghua University, Shenzhen, China

<sup>2</sup>PCL Research Center of Networks and Communications, Peng Cheng Laboratory, Shenzhen, China

<sup>3</sup>Department of Computing, The Hong Kong Polytechnic University, Hong Kong, China

<sup>4</sup>DAMO Academy, Alibaba Group

{dait14, zhang.yongbing, xiast}@sz.tsinghua.edu.cn, {csjcai, cslzhang}@comp.polyu.edu.hk

#### Abstract

Recently, deep convolutional neural networks (CNNs) have been widely explored in single image super-resolution (SISR) and obtained remarkable performance. However, most of the existing CNN-based SISR methods mainly focus

is ill-posed since multiple HR solutions can map to any LR input. Therefore, a great number of SR methods have been proposed, ranging from early interpolation-based [37] and model-based [4], to recent learning-based methods [32, 39].

The early developed interpolated-based methods (e.g., bilinear and bicubic methods) are simple and efficient but

#### Second-order attention network for single image super-resolution

T Dai, J Cai, Y Zhang, ST Xia... - Proceedings of the IEEE ..., 2019 - openaccess.thecvf.com

Recently, deep convolutional neural networks (CNNs) have been widely explored in single image super-resolution (SISR) and obtained remarkable performance. However, most of the existing CNN-based SISR methods mainly focus on wider or deeper architecture design, neglecting to explore the feature correlations of intermediate layers, hence hindering the representational power of CNNs. To address this issue, in this paper, we propose a second-order attention network (SAN) for more powerful feature expression and feature correlation ...

☆ 275회 인용 관련 학술자료 전체 8개의 버전

# SAN

## SAN Framework

- ❖ Feature간 관계 정보를 활용할 수 있는 모듈을 활용
  - ✓ Local-source residual attention group (LSRAG) module
- ❖ Residual Block, Channel Attention, Non-local block 활용

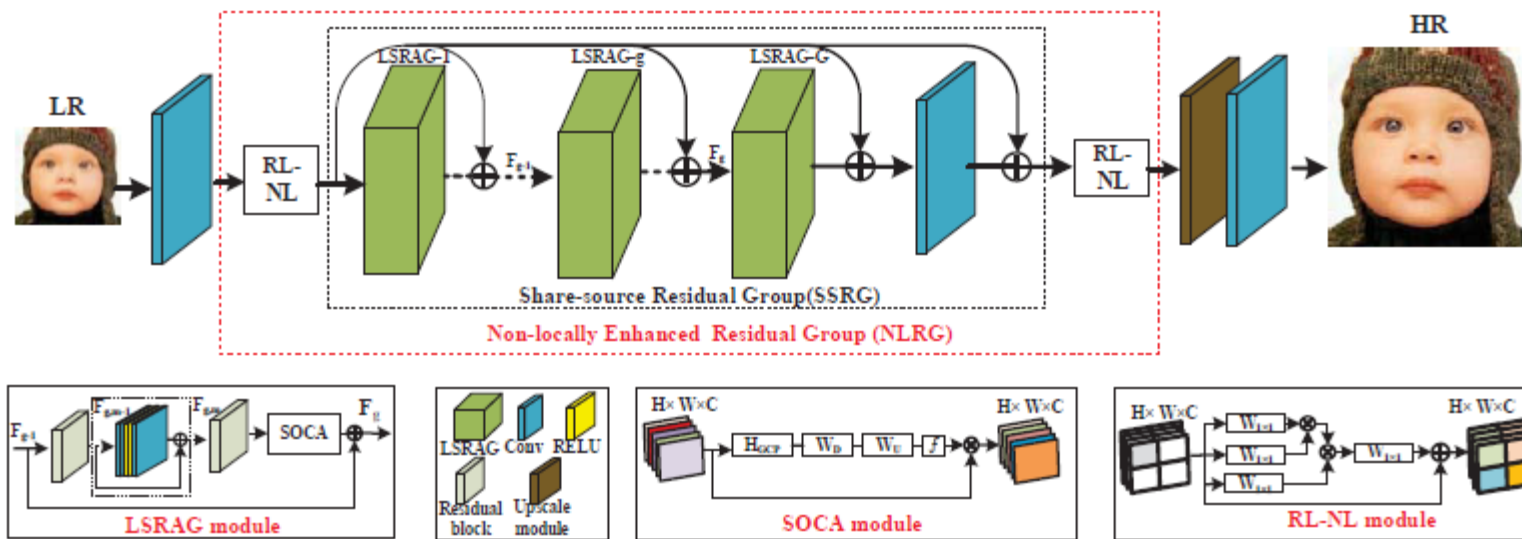
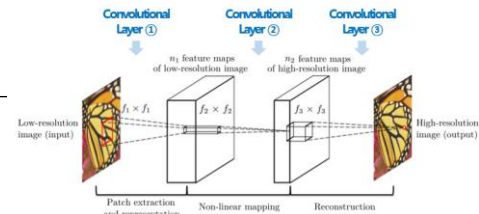
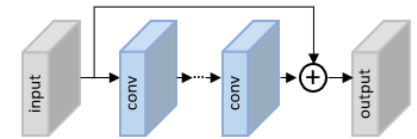


Figure 2. Framework of the proposed second-order attention network (SAN) and its sub-modules.

출처: Dai, T., Cai, J., Zhang, Y., Xia, S. T., & Zhang, L. (2019). Second-order attention network for single image super-resolution. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 11065-11074).

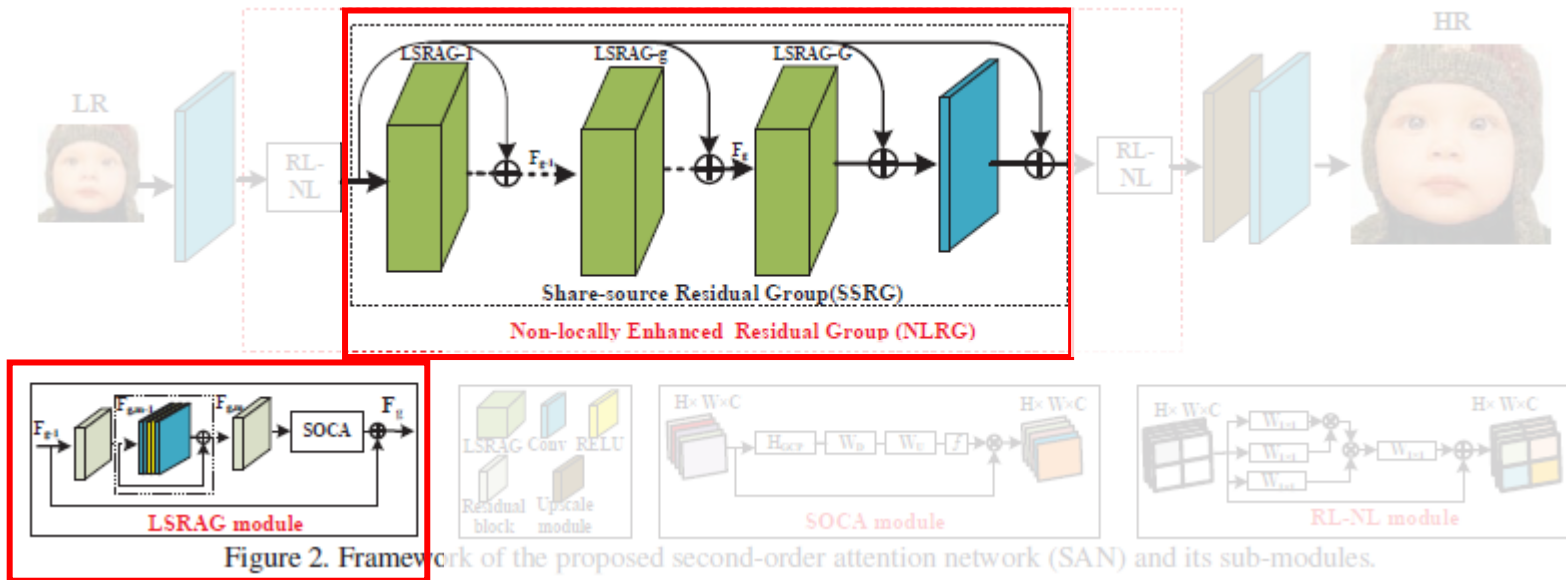
# SAN

① Residual Learning



## SAN Framework – LSRAG module

- ❖ LSRAG 모듈 안과 밖에 입력 feature 값을 더하는 Residual learning 존재
- ❖ 입력 feature가 지나는 정보를 계속해서 활용하며 학습을 진행



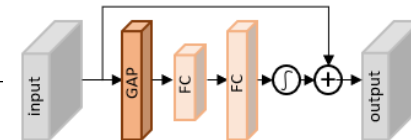
Residual Block: 입력 feature 정보를 활용하며 학습을 진행

출처: Dai, T., Cai, J., Zhang, Y., Xia, S. T., & Zhang, L. (2019). Second-order attention network for single image super-resolution. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 11065-11074).



# SAN

③ Channel attention



## SAN Framework – SOCA module

- ❖ LSRAG 모듈 마지막에 Second order Attention Network (SOCA) 모듈 존재
- ❖ Channel Attention을 활용해 Channel 간 관계 정보 고려해 학습 진행

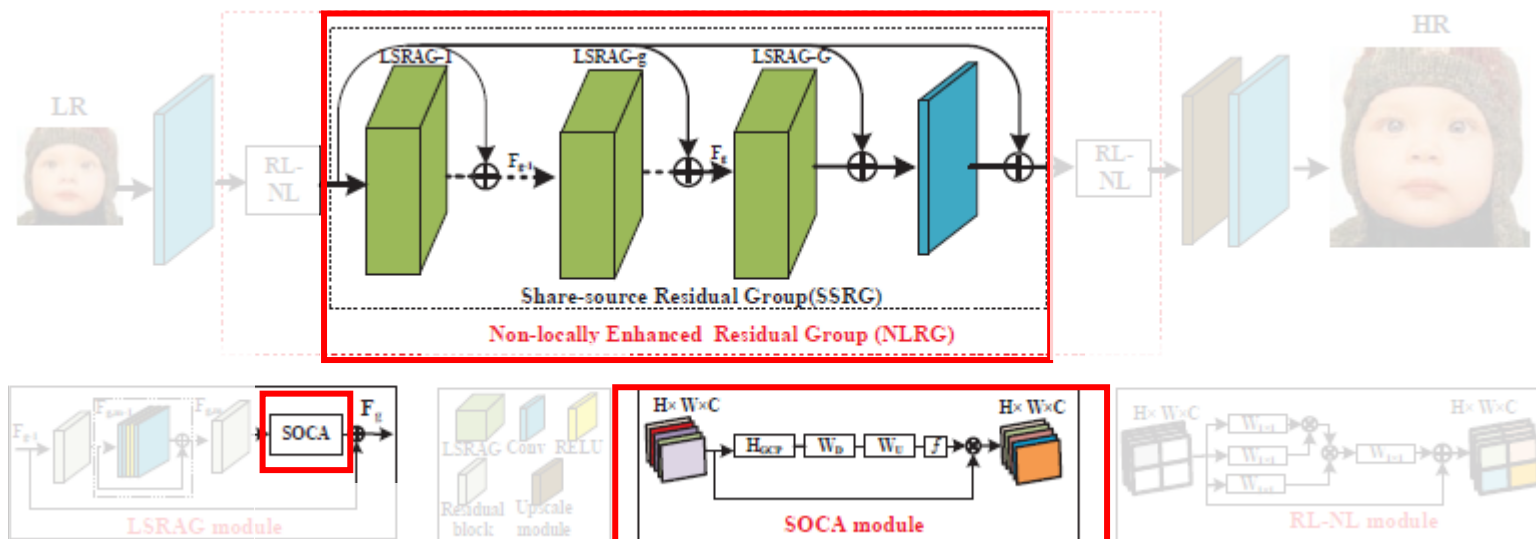


Figure 2. Framework of the proposed second-order attention network (SAN) and its sub-modules.

→ Channel Attention: Channel간 feature의 관계 정보를 고려해 학습

# SAN

## SAN Framework – RL-NL

- ❖ 전체 네트워크 앞과 마지막 부분에 Region-level non-local (RL-NL) 모듈 적용
- ❖ Feature를 특정 수의 지역으로 나눠 pixel간 관계 정보를 고려해 학습

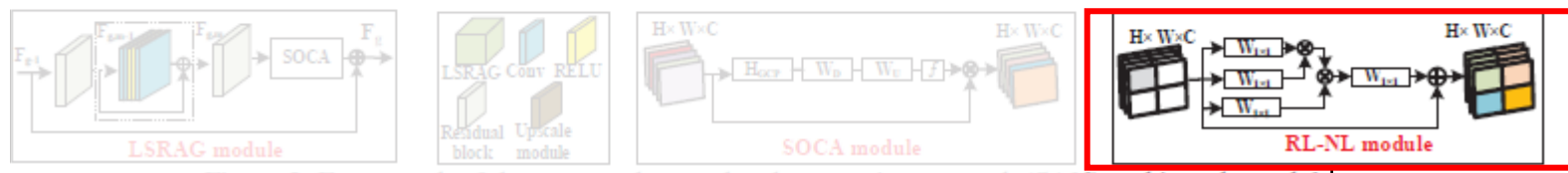
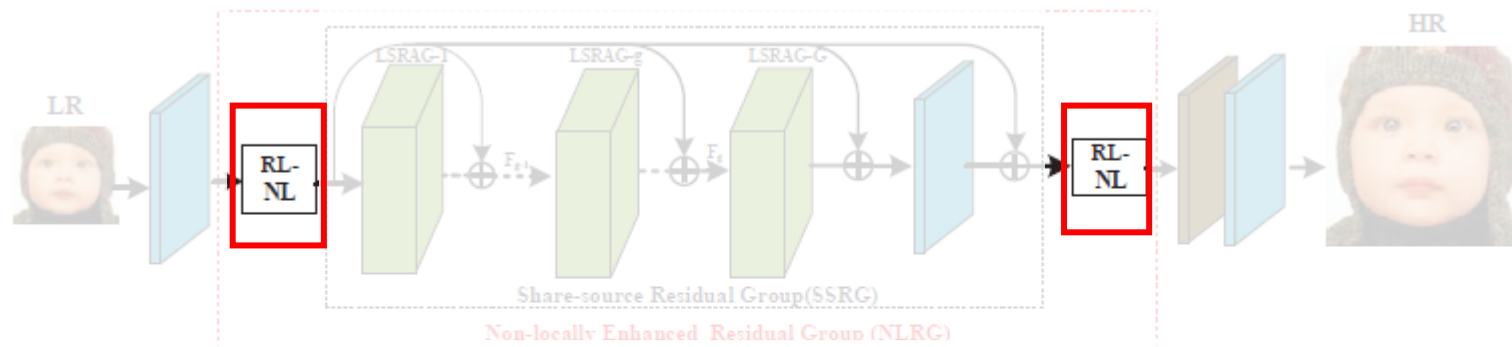
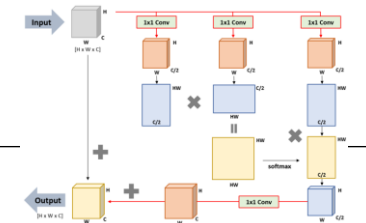


Figure 2. Framework of the proposed second-order attention network (SAN) and its sub-modules.

→ Non-local block – feature 내 pixel간 관계 정보를 고려해 학습

출처: Dai, T., Cai, J., Zhang, Y., Xia, S. T., & Zhang, L. (2019). Second-order attention network for single image super-resolution. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 11065-11074).

# SAN

## SAN Framework

- ❖ Feature간 관계 정보를 활용할 수 있는 모듈을 활용
  - ✓ Local-source residual attention group (LSRAG) module
- ❖ Residual Block, Channel Attention, Non-local block 활용

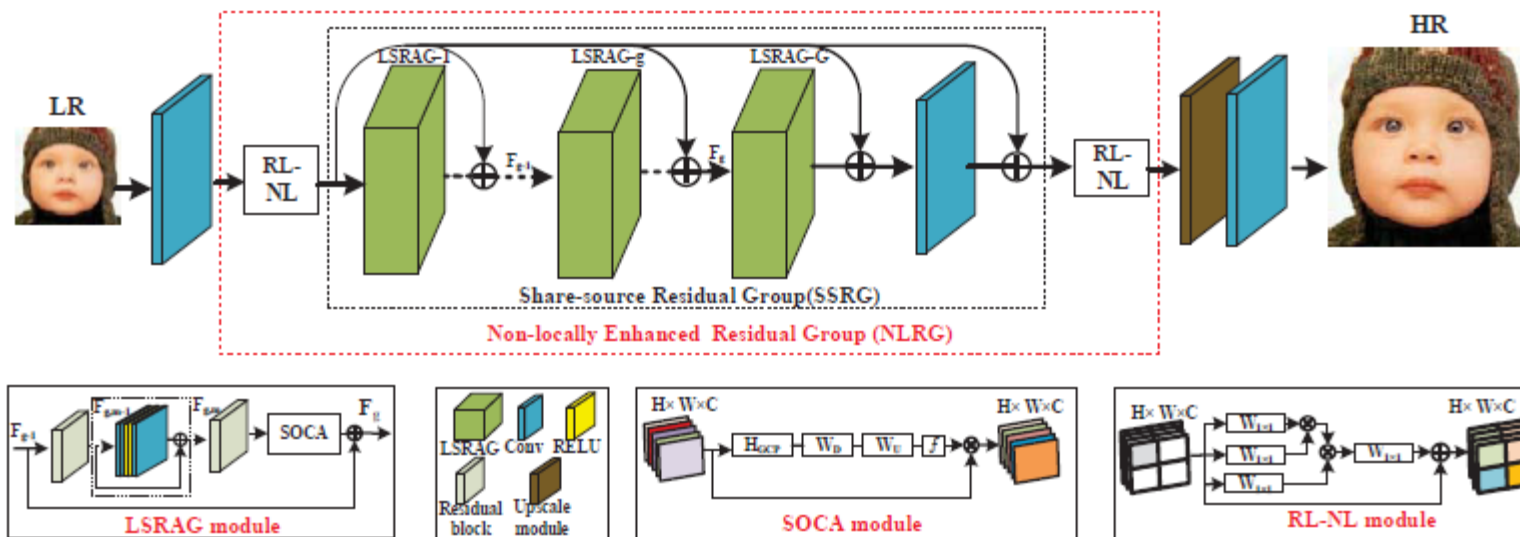
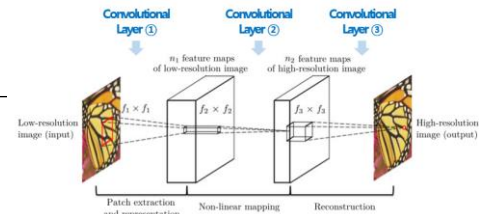


Figure 2. Framework of the proposed second-order attention network (SAN) and its sub-modules.

출처: Dai, T., Cai, J., Zhang, Y., Xia, S. T., & Zhang, L. (2019). Second-order attention network for single image super-resolution. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 11065-11074).

## SAN 결과 - 정량 평가

- ❖ 총 5개 Test 데이터셋과 4개 이미지 복원 비율에 대해 비교·평가 진행
- ❖ 대부분 결과에서 제안한 SAN 모델 성능이 가장 좋았음

Table 2. Quantitative results with BI degradation model.

Method		Set5	Set14	BSD100	Urban100	Manga109							
		PSNR/ SSIM	PSNR/ SSIM	PSNR/ SSIM	PSNR/ SSIM	PSNR/ SSIM							
Bicubic	2	33.66/9299	30.24/8688	29.56/8431	26.88/8403	30.80/9339	Bicubic	4	28.42/8104	26.00/7027	25.96/6675	23.14/6577	24.89/7866
SRCNN	2	36.66/9542	32.45/9067	31.36/8879	29.50/8946	35.60/9663	SRCNN	4	30.48/8628	27.50/7513	26.90/7101	24.52/7221	27.58/8555
FSRCNN	2	37.05/9560	32.66/9090	31.53/8920	29.88/9020	36.67/9710	FSRCNN	4	30.72/8660	27.61/7550	26.98/7150	24.62/7280	27.90/8610
VDSR	2	37.53/9590	33.05/9130	31.90/8960	30.77/9140	37.22/9750	VDSR	4	31.35/8830	28.02/7680	27.29/7276	25.18/7540	28.83/8870
LapSRN	2	37.52/9591	33.08/9130	31.08/8950	30.41/9101	37.27/9740	LapSRN	4	31.54/8850	28.19/7720	27.32/7270	25.21/7560	29.09/8900
MemNet	2	37.78/9597	33.28/9142	32.08/8978	31.31/9195	37.72/9740	MemNet	4	31.74/8893	28.26/7723	27.40/7281	25.50/7630	29.42/8942
EDSR	2	38.11/9602	33.92/9195	32.32/9013	32.93/9351	39.10/9773	EDSR	4	32.46/8968	28.80/7876	27.71/7420	26.64/8033	31.02/9148
SRMD	2	37.79/9601	33.32/9159	32.05/8985	31.33/9204	38.07/9761	SRMD	4	31.96/8925	28.35/7787	27.49/7337	25.68/7731	30.09/9024
NLRN	2	38.00/9603	33.46/9159	32.19/8992	31.81/9246	—/—	DBPN	4	32.47/8980	28.82/7860	27.72/7400	26.38/7946	30.91/9137
DBPN	2	38.09/9600	33.85/9190	32.27/9000	32.55/9324	38.89/9775	NLRG	4	31.92/8916	28.36/7745	27.48/7346	25.79/7729	—/—
RDN	2	38.24/9614	34.01/9212	32.34/9017	32.89/9353	39.18/9780	RDN	4	32.47/8990	28.81/7871	27.72/7419	26.61/8028	31.00/9151
RCAN	2	38.27/9614	34.11/9216	32.41/9026	33.34/9384	39.43/9786	RCAN	4	32.62/9001	28.86/7888	27.76/7435	26.82/8087	31.21/9172
SAN	2	38.31/9620	34.07/9213	32.42/9028	33.10/9370	39.32/9792	SAN	4	32.64/9003	28.92/7888	27.78/7436	26.79/8068	31.18/9169
SAN+	2	<b>38.35/9619</b>	<b>34.44/9244</b>	<b>32.50/9038</b>	<b>33.73/9416</b>	<b>39.72/9797</b>	SAN+	4	<b>32.70/9013</b>	<b>29.05/7921</b>	<b>27.86/7457</b>	<b>27.23/8169</b>	<b>31.66/9222</b>
Bicubic	3	30.39/8682	27.55/7742	27.21/7385	24.46/7349	26.95/8556	Bicubic	8	24.40/6580	23.10/5660	23.67/5480	20.74/5160	21.47/6500
SRCNN	3	32.75/9090	29.30/8215	28.41/7863	26.24/7989	30.48/9117	SRCNN	8	25.33/6900	23.76/5910	24.13/5660	21.29/5440	22.46/6950
FSRCNN	3	33.18/9140	29.37/8240	28.53/7910	26.43/8080	31.10/9210	FSRCNN	8	20.13/5520	19.75/4820	24.21/5680	21.32/5380	22.39/6730
VDSR	3	33.67/9210	29.78/8320	28.83/7990	27.14/8290	32.01/9340	SCN	8	25.59/7071	24.02/6028	24.30/5698	21.52/5571	22.68/6963
LapSRN	3	33.82/9227	29.87/8320	28.82/7980	27.07/8280	32.21/9350	VDSR	8	25.93/7240	24.26/6140	24.49/5830	21.70/5710	23.16/7250
MemNet	3	34.09/9248	30.01/8350	28.96/8001	27.56/8376	32.51/9369	LapSRN	8	26.15/7380	24.35/6200	24.54/5860	21.81/5810	23.39/7350
EDSR	3	34.65/9280	3.52/8462	29.25/8093	28.80/8653	34.17/9476	MemNet	8	26.16/7414	24.38/6199	24.58/5842	21.89/5825	23.56/7387
SRMD	3	34.12/9254	30.04/8382	28.97/8025	27.57/8398	33.00/9403	MSLap	8	26.34/7558	24.57/6273	24.65/5895	22.06/5963	23.90/7564
NLRG	3	34.27/9266	30.16/8374	29.06/8026	27.93/8453	—/—	EDSR	8	26.96/7762	24.91/6420	24.81/5985	22.51/6221	24.69/7841
RDN	3	34.71/9296	30.57/8468	29.26/8093	28.80/8653	34.13/9484	DBPN	8	27.21/7840	25.13/6480	24.88/6010	22.73/6312	25.14/7987
RCAN	3	34.74/9299	30.64/8481	29.32/8111	29.08/8702	34.43/9498	SAN	8	27.22/7829	25.14/6476	24.88/6011	22.70/6314	24.85/7906
SAN	3	34.75/9300	30.59/8476	29.33/8112	28.93/8671	34.30/9494	SAN+	8	<b>27.30/7849</b>	<b>25.23/6493</b>	<b>24.97/6031</b>	<b>22.91/6369</b>	<b>25.17/7964</b>
SAN+	3	<b>34.89/9306</b>	<b>30.77/8498</b>	<b>29.38/8121</b>	<b>29.29/8730</b>	<b>34.74/9512</b>							

출처: Dai, T., Cai, J., Zhang, Y., Xia, S. T., & Zhang, L. (2019). Second-order attention network for single image super-resolution. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 11065-11074).

## SAN 결과 - 정성 평가

- ❖ SAN 모델 적용시 HR 이미지에 가장 가깝게 이미지가 복원되었음을 확인

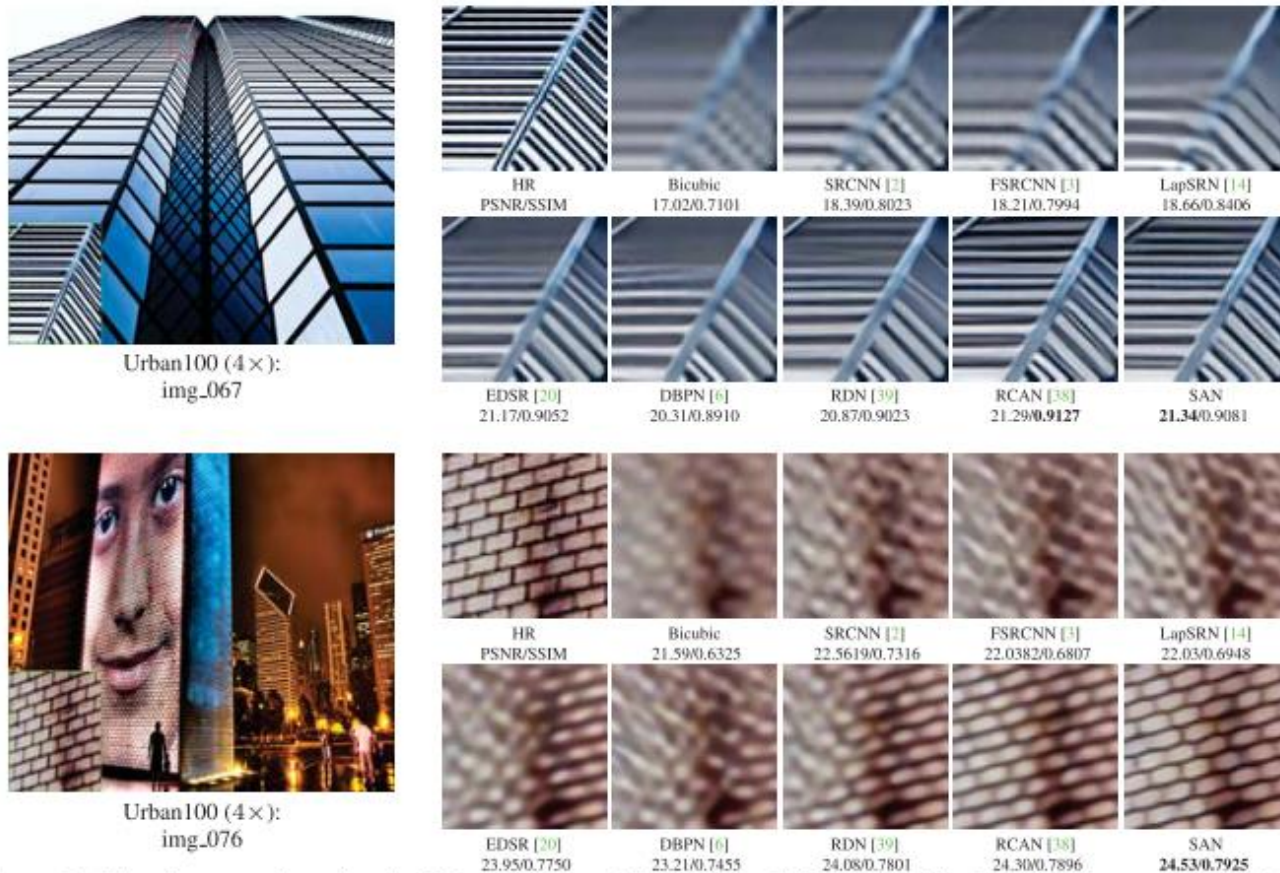


Figure 3. Visual comparison for 4× SR with BI model on Urban100 dataset. The best results are **highlighted**

출처: Dai, T., Cai, J., Zhang, Y., Xia, S. T., & Zhang, L. (2019). Second-order attention network for single image super-resolution. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 11065-11074).

---

# 5. Conclusion

---

# Conclusion

---

## ❖ Conclusion

- ① Image Super-Resolution 문제를 풀기 위해서는 저해상도 이미지 크기와 고해상도 이미지 크기를 잘 맞춰주는 것이 선행되어야 함
- ② 일반적인 딥러닝 모델에서 사용하는 모든 방법들을 적용할 수 있음
- ③ 고해상도 변환이 잘되었음을 판단하기 위해서는 PSNR, SSIM과 같은 지표가 필요
- ④ 평가에서는 정량적인 지표에 더하여 정성적인 지표도 항상 같이 고려되어야 함
- ⑤ Data Augmentation, Self-Supervised Learning, Generative Adversarial Network 등 다양한 방법론을 적용해서 Image Super-Resolution 모델 성능 향상을 기대

---

# 감사합니다.

---



## ❖ 출처

- Wang, Z., Chen, J., & Hoi, S. C. (2020). Deep learning for image super-resolution: A survey. *IEEE transactions on pattern analysis and machine intelligence*.
- Dong, C., Loy, C. C., He, K., & Tang, X. (2015). Image super-resolution using deep convolutional networks. *IEEE transactions on pattern analysis and machine intelligence*, 38(2), 295-307.
- Dai, T., Cai, J., Zhang, Y., Xia, S. T., & Zhang, L. (2019). Second-order attention network for single image super-resolution. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 11065-11074).
- <https://www.yna.co.kr/view/AKR20191108058400063>
- <http://www.digitaltoday.co.kr/news/articleView.html?idxno=233998>
- <https://m.blog.naver.com/PostView.nhn?blogId=dudcos101&logNo=221421989207&proxyReferer=https:%2F%2Fwww.google.com%2F>
- <https://github.com/MagmaTart/Paper-Reading/blob/master/summarys/SRGAN.md>
- <https://hoya012.github.io/blog/Single-Image-Super-Resolution-Overview/>
- <https://jaejunyoo.blogspot.com/2019/05/deep-learning-for-SISR-survey-1.html>
- <https://bskyvision.com/392>



---

# Appendix.

---