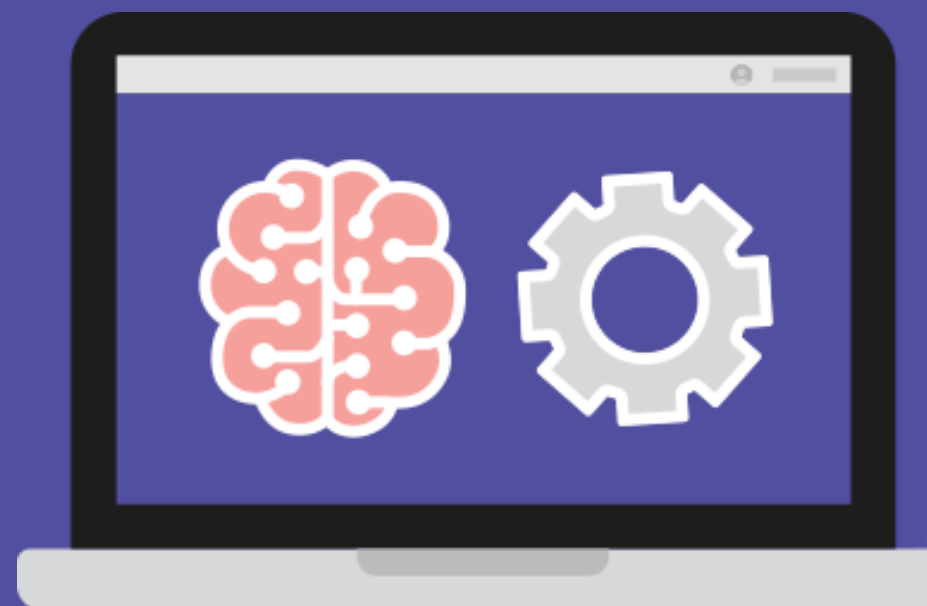


/* elice */

양재 AI School 인공지능 캠프

Lecture 19

CNN을 활용한 여러 신경망 모델



박상수 선생님

커리큘럼

1 ○

컨볼루션이 사용된 다양한 알고리즘

Deconvolution, Dilated 컨볼루션을 파악합니다.

2 ○

CNN을 다른 용도로 사용하는 신경망

물체를 분류하고 인식하는 Region-CNN과
이미지의 품질을 높이는 Super Resolution CNN을 배웁니다.

목차

1. 컨볼루션이 사용된 알고리즘
2. ZFNet
3. Region CNN
4. Super Resolution CNN
5. GAN

컨볼루션이 사용된 알고리즘

컨볼루션 기반의 변형 알고리즘

Deconvolution

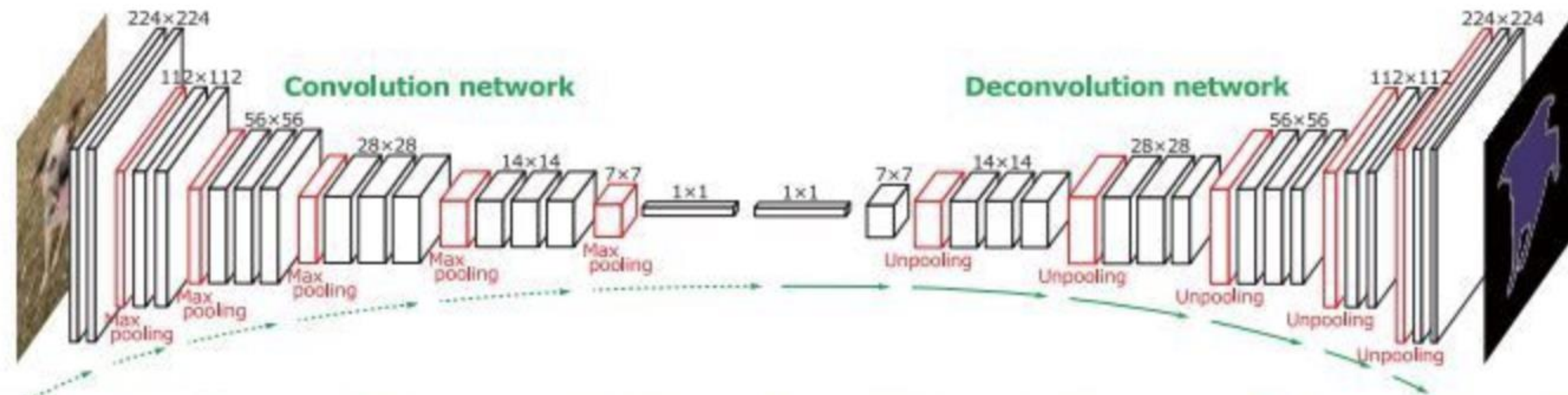
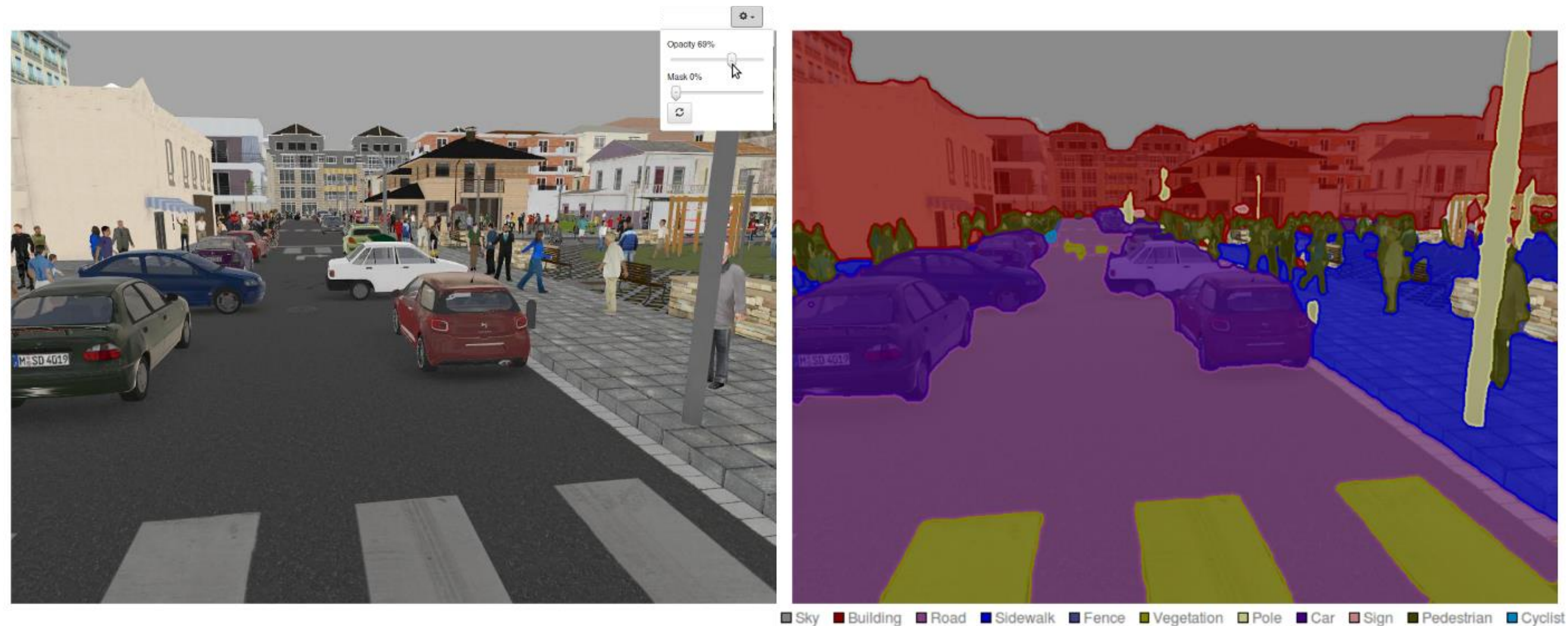


Figure 2. Overall architecture of the proposed network. On top of the convolution network based on VGG 16-layer net, we put a multi-layer deconvolution network to generate the accurate segmentation map of an input proposal. Given a feature representation obtained from the convolution network, dense pixel-wise class prediction map is constructed through multiple series of unpooling, deconvolution and rectification operations.

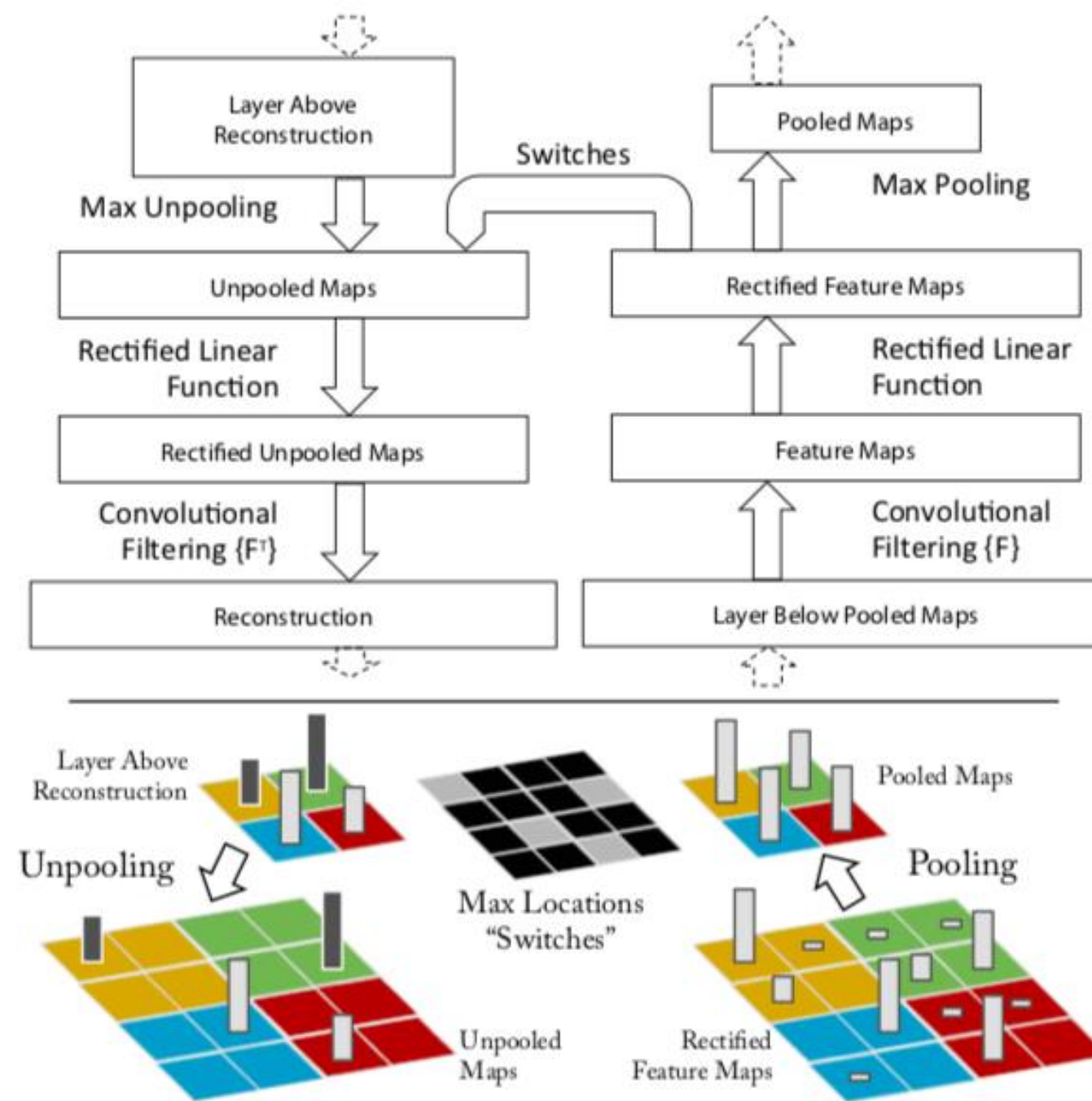
컨볼루션이 된 이미지를 원래로 돌리는 과정
(Semantic Segmentation 분야에서 사용됨)

Semantic Segmentation



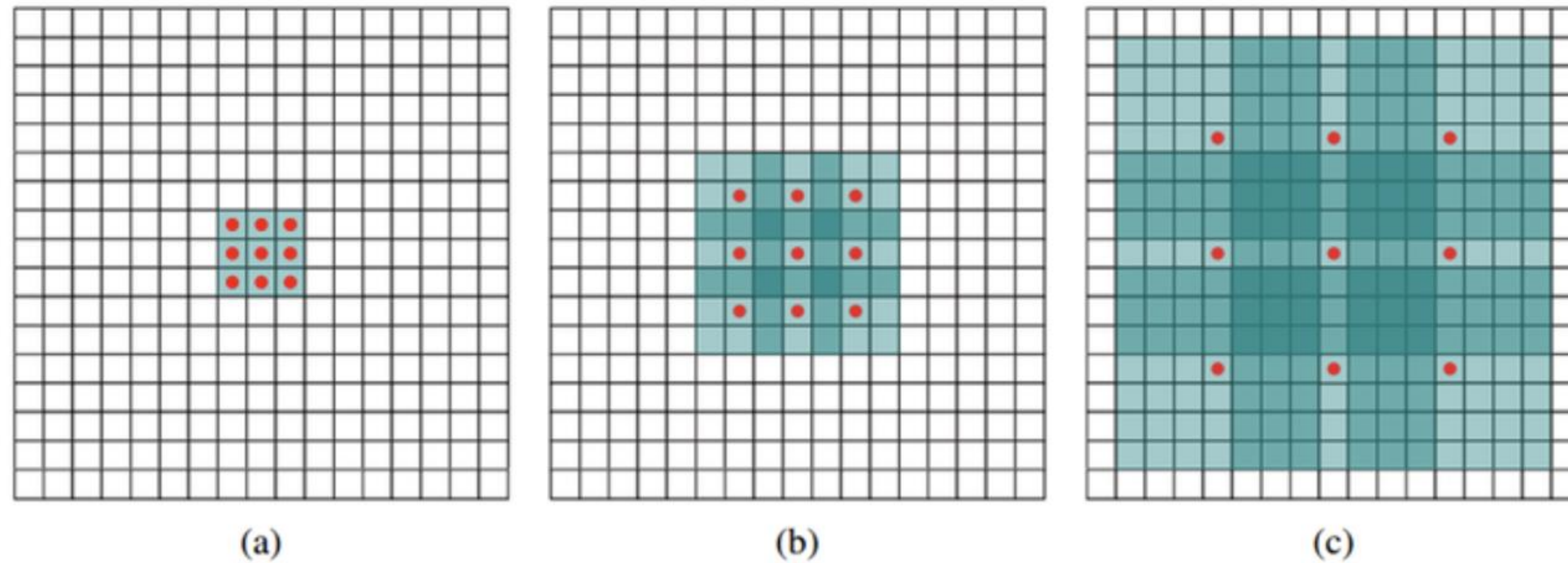
사진을 분류하는 것에 그치지 않고, 그 장면을
완벽하게 이해해야하는 높은 수준의 문제

Deconvolution



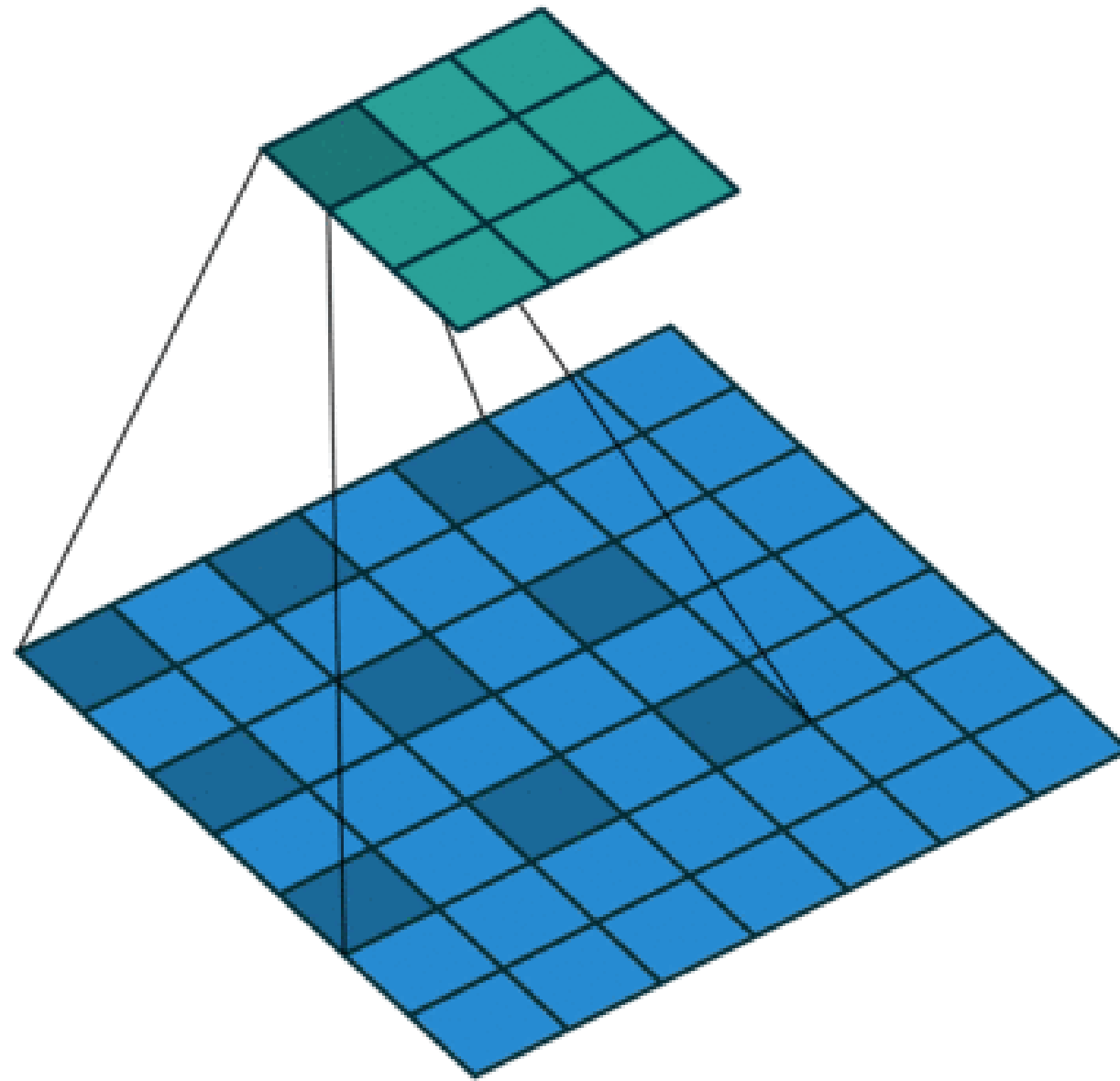
ZFNet에서 사용한 시각화와 동일한 방법

Dilated Convolution



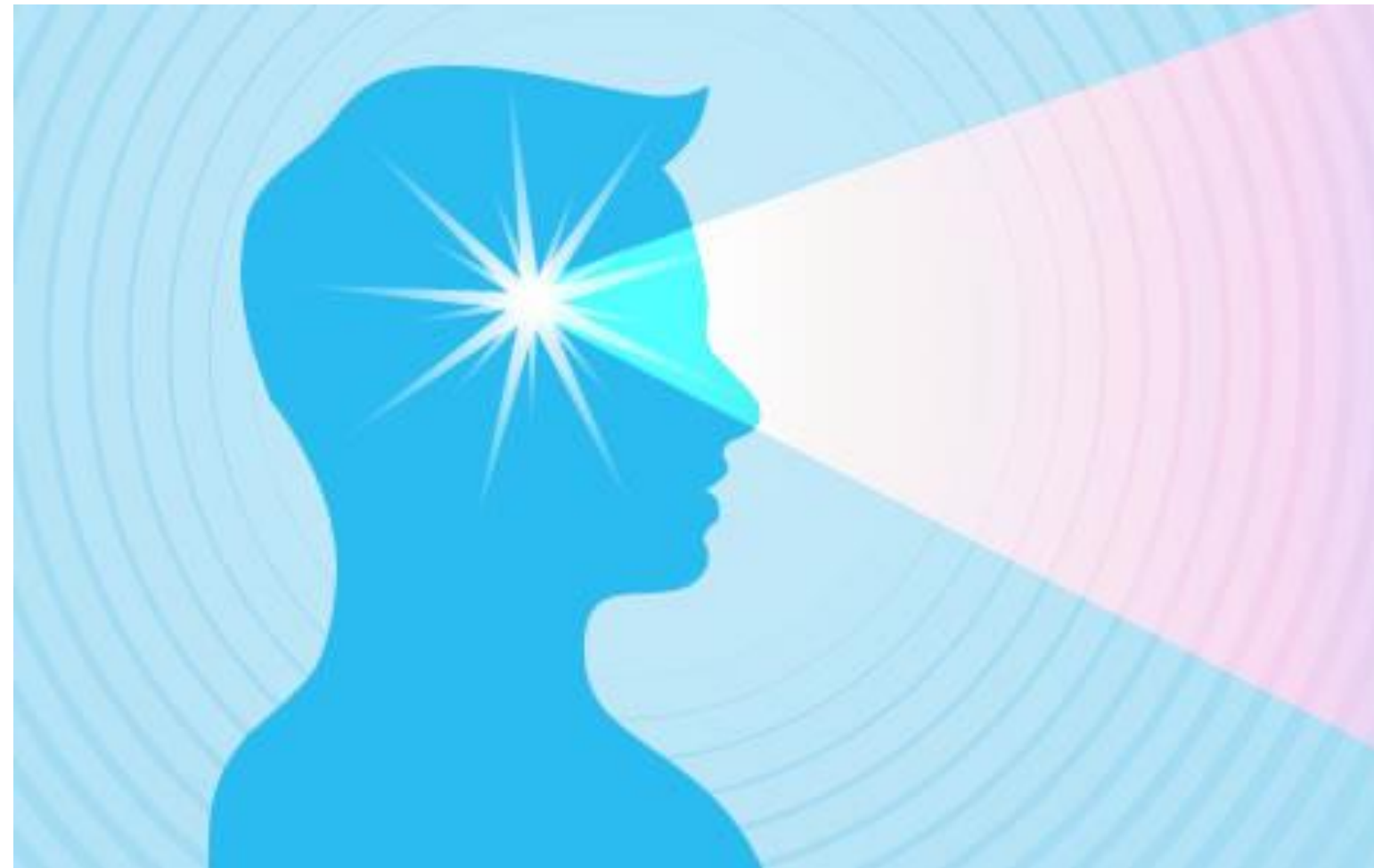
빨간색 점의 위치에 있는 픽셀만 사용하여 컨볼루션
(Atrous Convolution, Atrous는 구멍을 의미함)
빨간색이 아닌 부분은 모두 0값으로 채움

Dilated Convolution



Real-Time Segmentation 분야에서 주로 사용
넓은 시야가 필요하고 여러 Convolution이나 큰 커널을 사용할 여유가
없는 경우에 사용

Visualizing

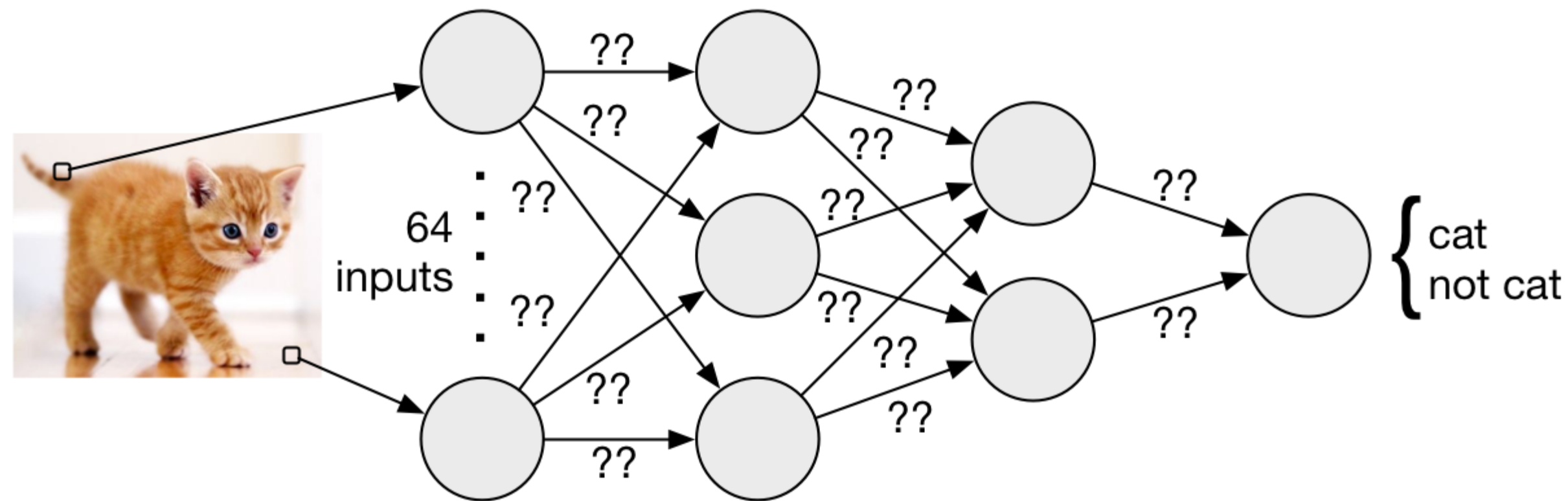


시각화 (Visualizing) 기법을 사용하여 해결

ZFNet

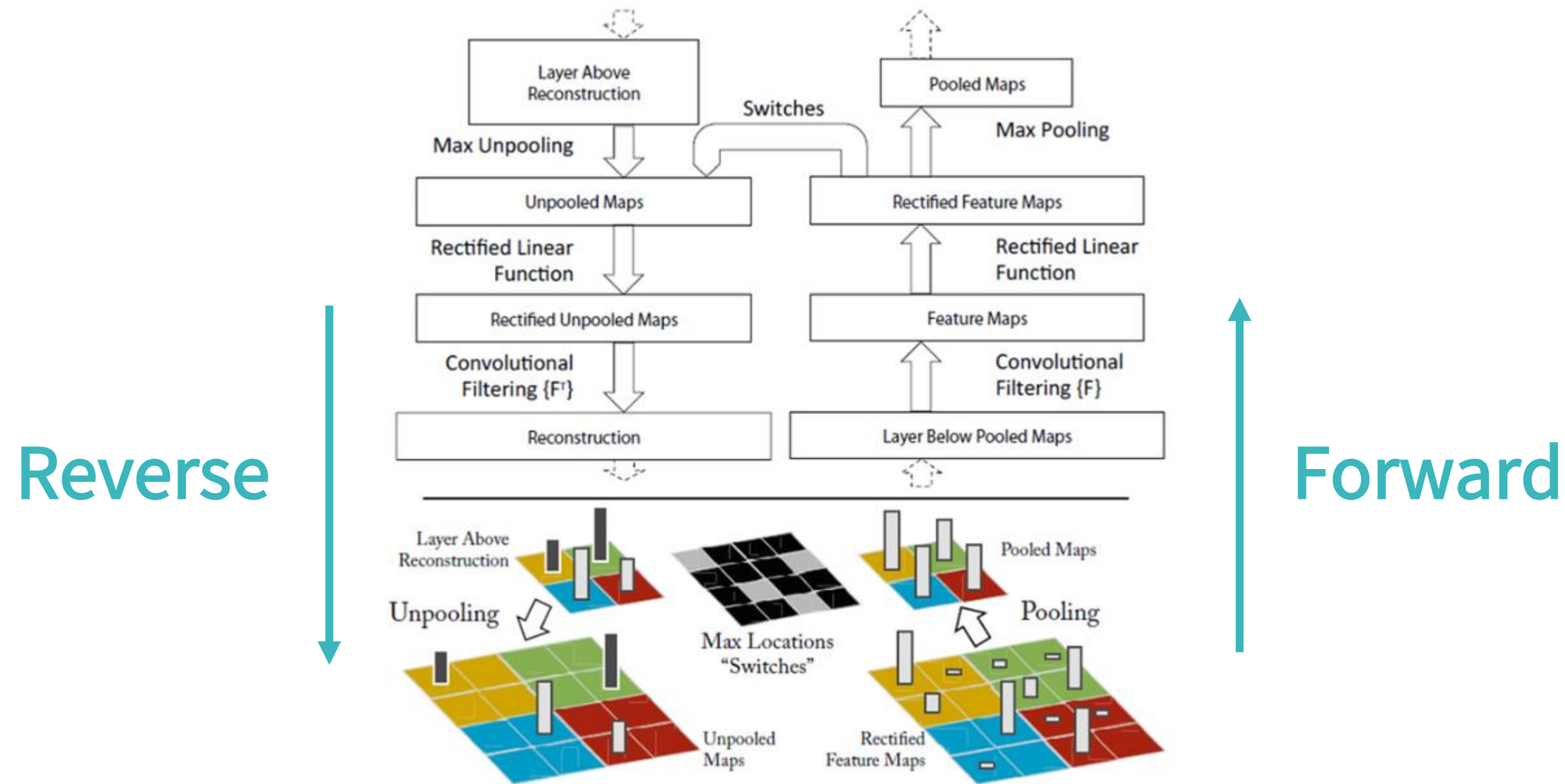
학습 되는 것을 시각적으로 볼 수는 없을까 ?

Visualizing & Understanding CNN



CNN을 학습하는 과정에서 무슨 원리로 좋은 성능을 얻는지
Hyper parameter를 어떻게 설정할 것인지 등등
개발한 신경망의 구조가 최적인지 확인하는 것은 어려움

Reverse construction



Forward 방향: Conv-ReLU-Pool

Reverse 방향: Pool-ReLU-Conv

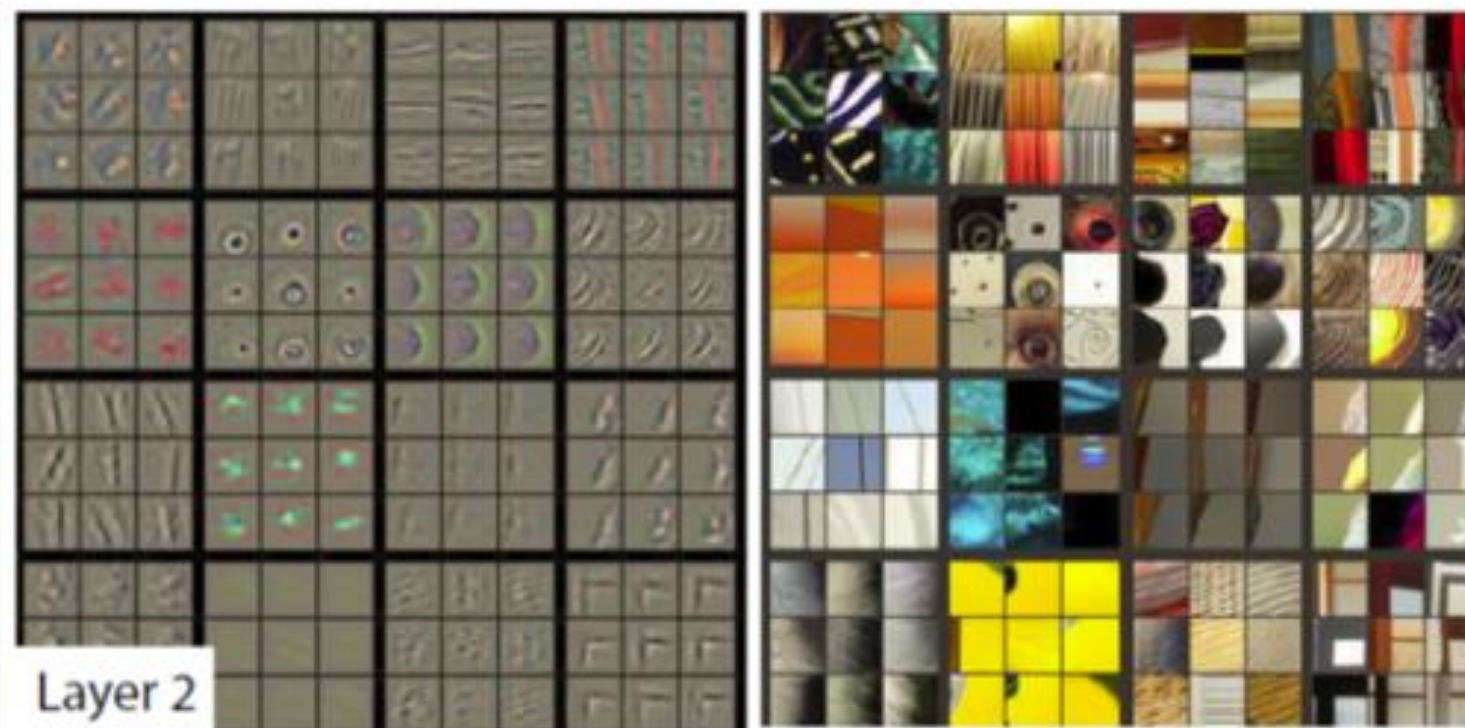
Max pooling 방법을 사용했는데, 어떻게 복구 ???
가장 큰 값의 위치를 기억 !

Feature visualization



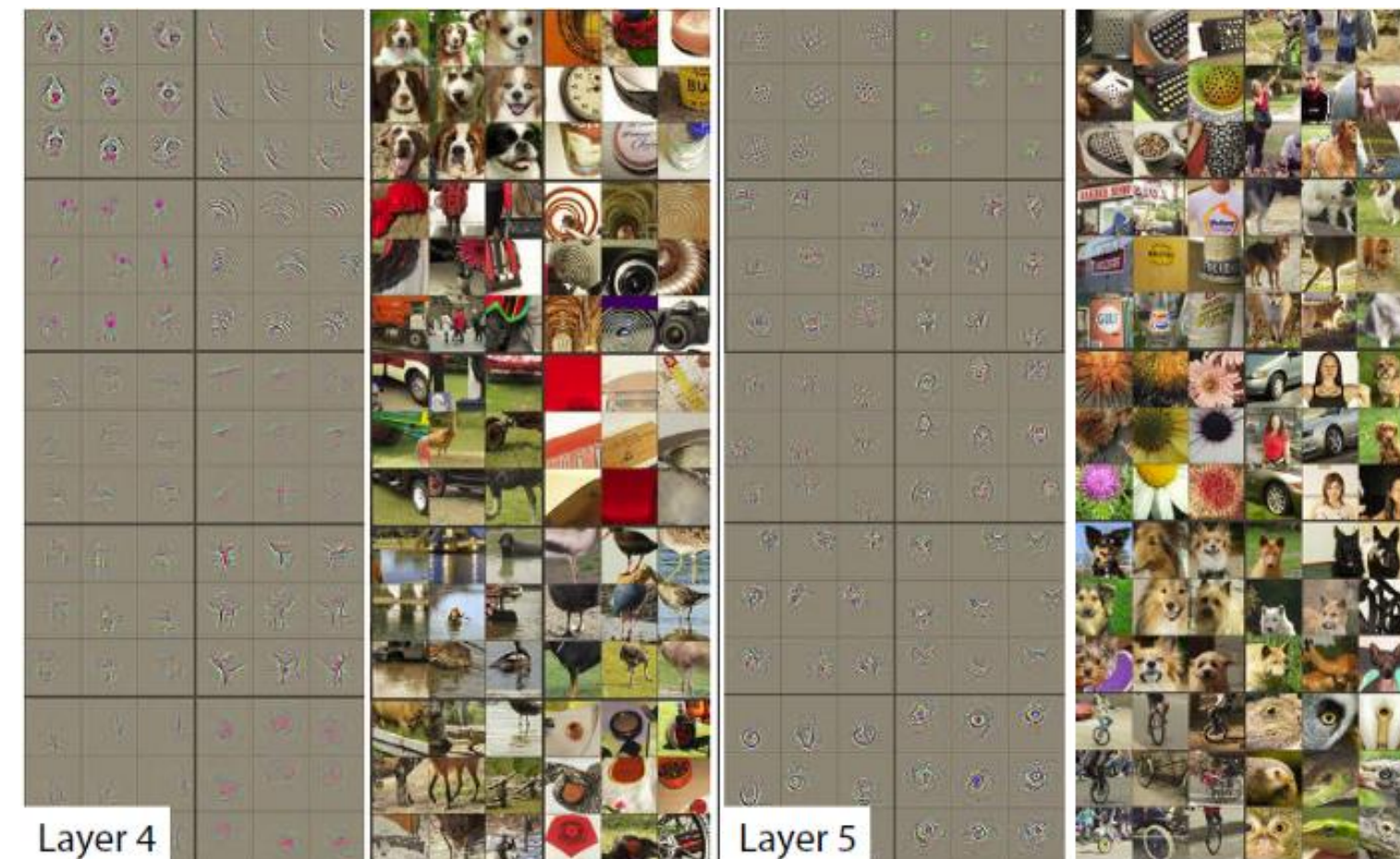
Layer 1

Extreme low and high frequencies.
No mid frequencies.



Layer 2

Aliasing problem



Layer 4

Layer 5

Region CNN

물체가 위치한 정보도 찾는 신경망

Classification and Detection

**Classification
+ Localization**



CAT

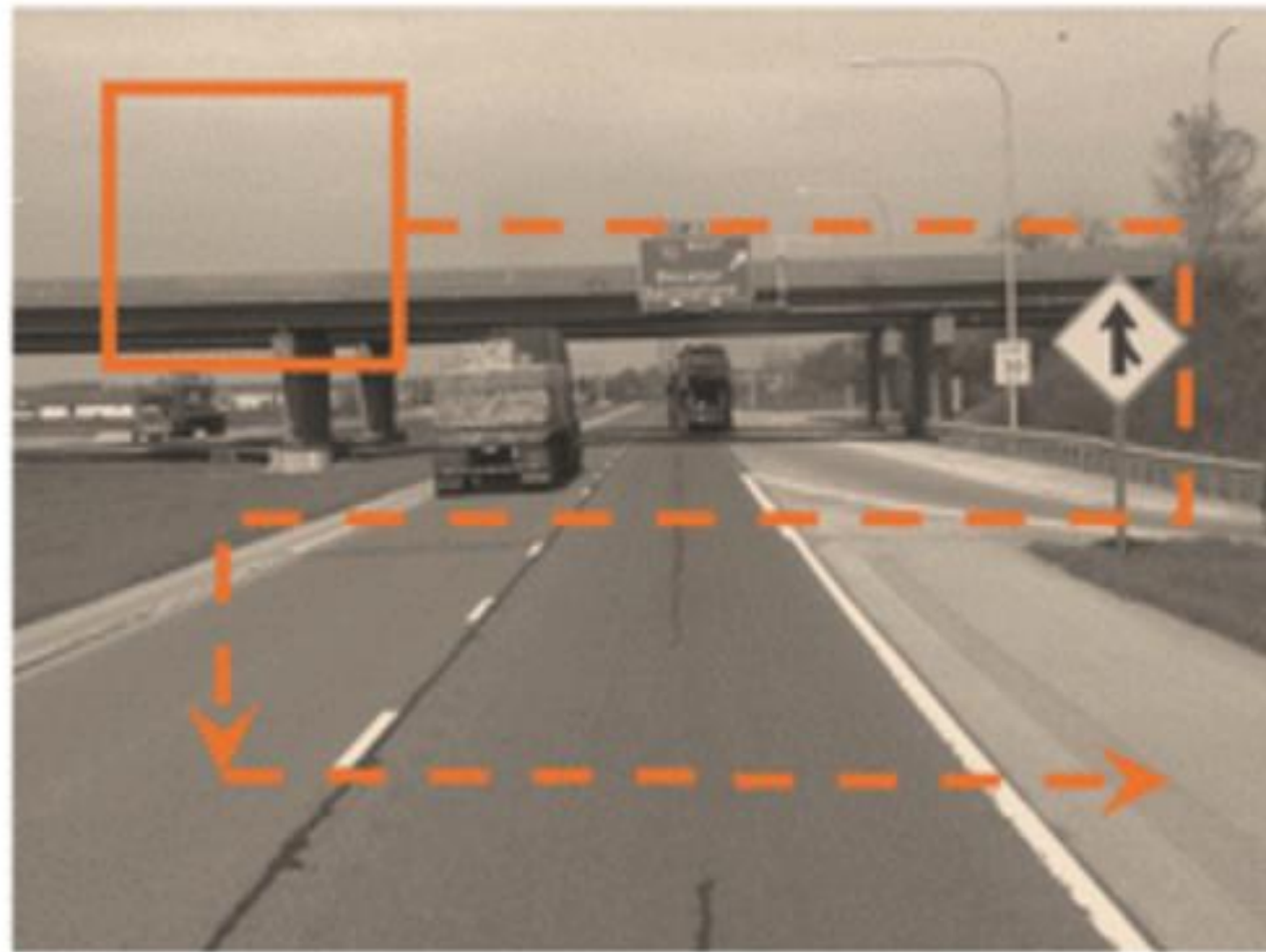
**Object
Detection**



DOG, DOG, CAT

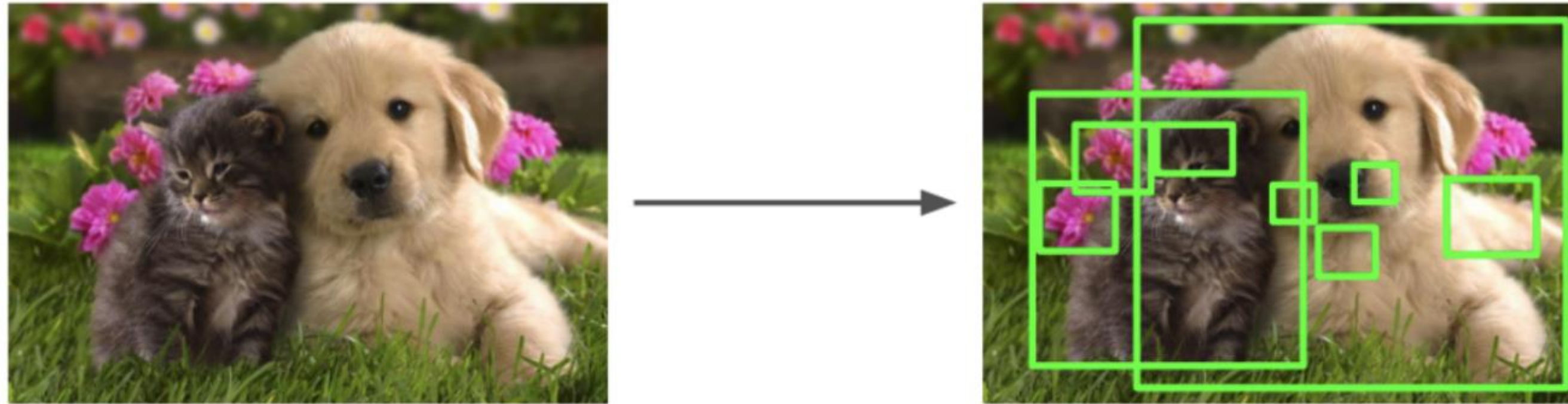
Classification: 물체의 이름을 맞추는 것
Detection: 물체의 이름을 맞추고, 위치를 찾는 것

Naïve Approach



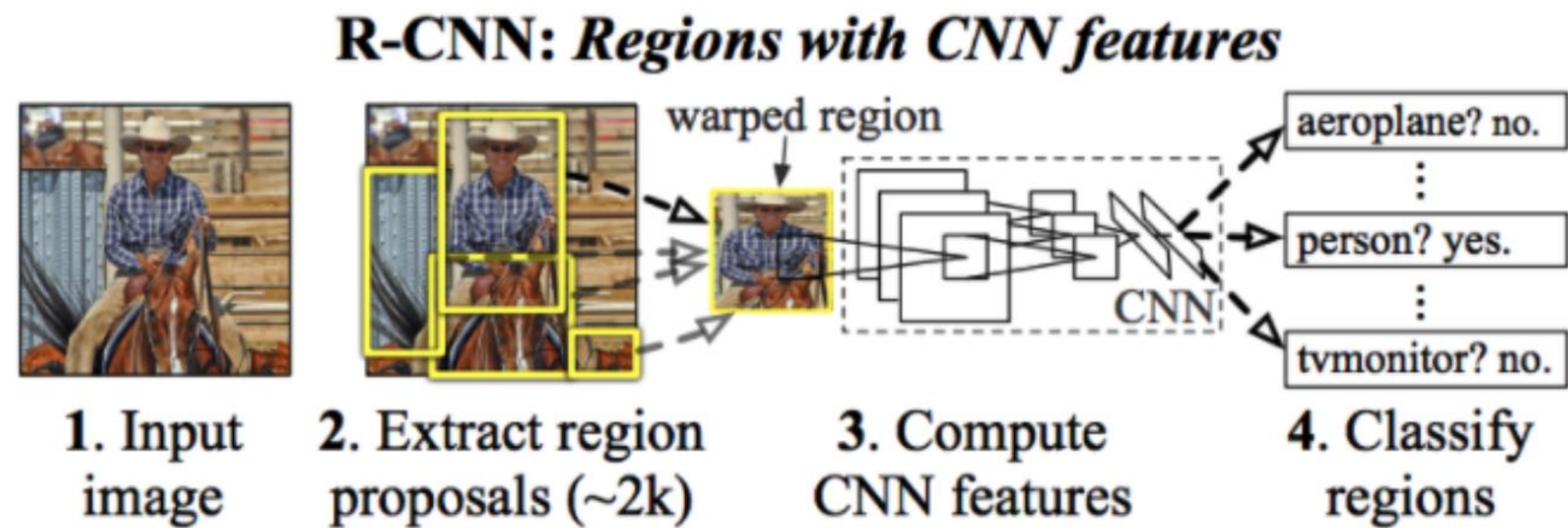
모든 크기의 영역에 대해 **Sliding Window** 방식으로
이미지를 모두 탐색하면서 분류하는 방식
하지만, 탐색해야 하는 영역이 너무 많은 문제

Region Proposal



물체가 있을 법한 영역을 빠른 속도로 찾는 알고리즘
Sliding Window 방식에 비해 탐색 영역이 줄기 때문에
훨씬 빠른 속도로 물체 인식 가능

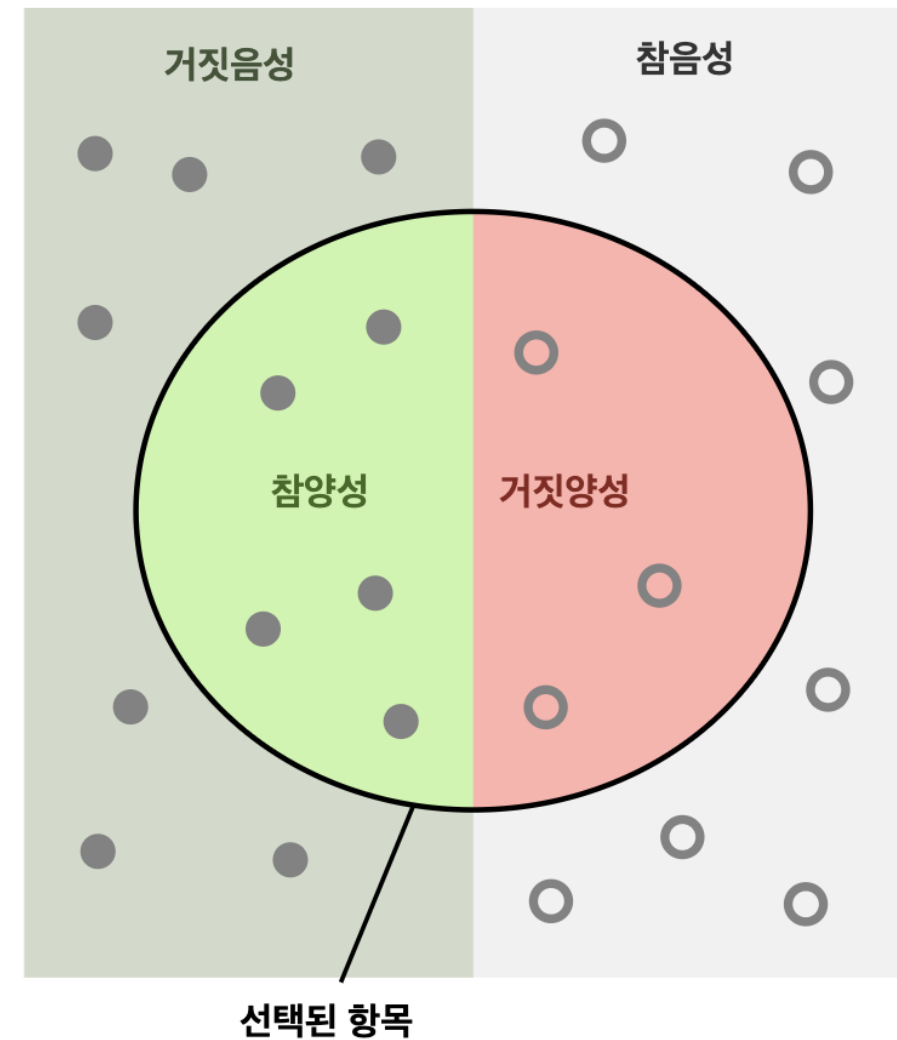
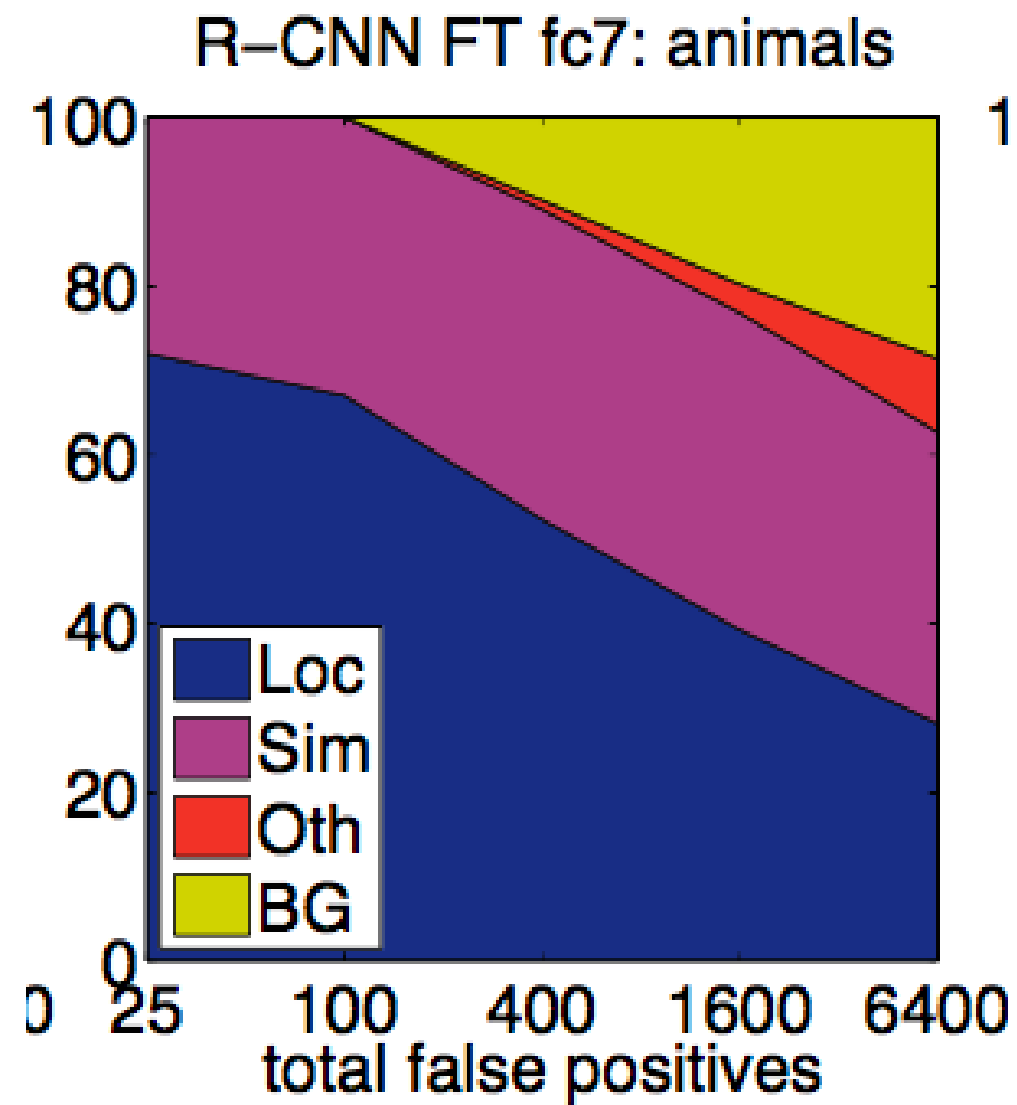
Region Proposal + CNN



물체가 존재할 영역에 대해 CNN을 적용

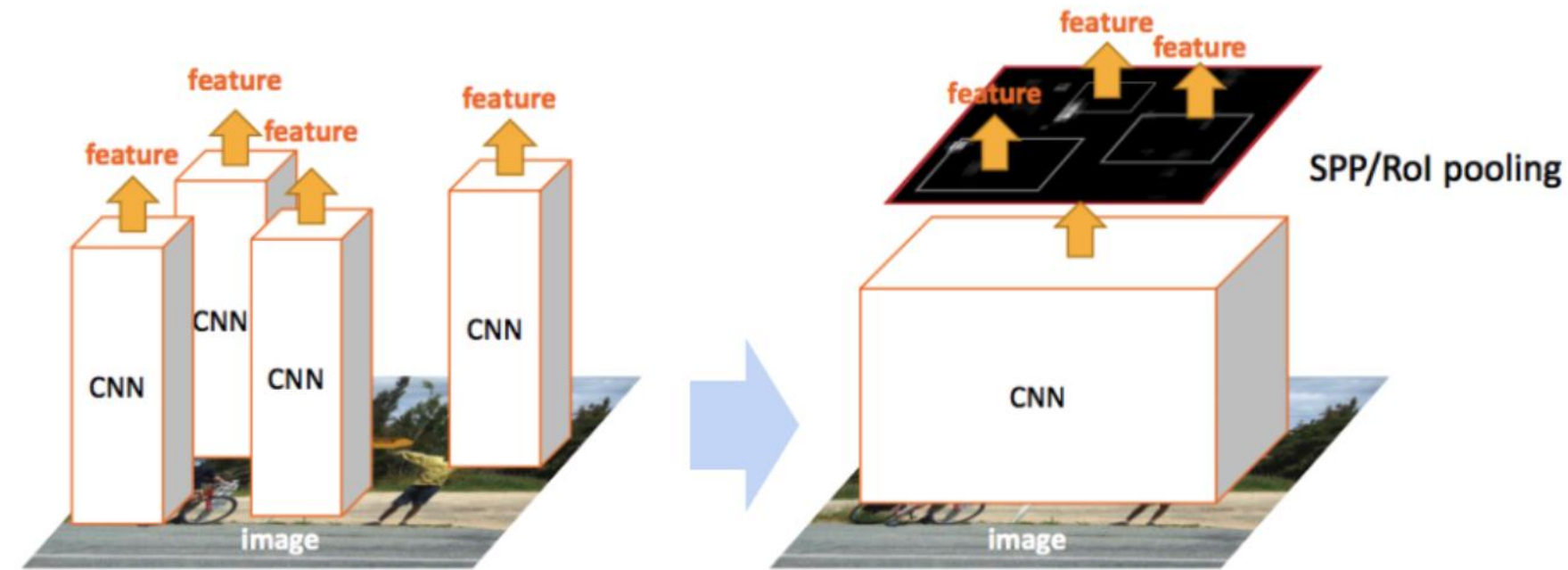
1. 이미지 입력
2. 이미지로부터 2,000개 정도의 Region Proposal 추출
3. Region Proposal을 CNN의 입력으로 하여 특징 추출
4. 분류 (SVM)

Region Proposal + CNN



Localization 취약성이 CNN이 Positional Invariance한 특징을 때문
Region Proposal 내에서 물체가 중앙이 아닌 곳에 있어도, CNN은
예측이 가능하기 때문에, 정확한 위치를 잡기에는 부족함

Fast R-CNN



R-CNN

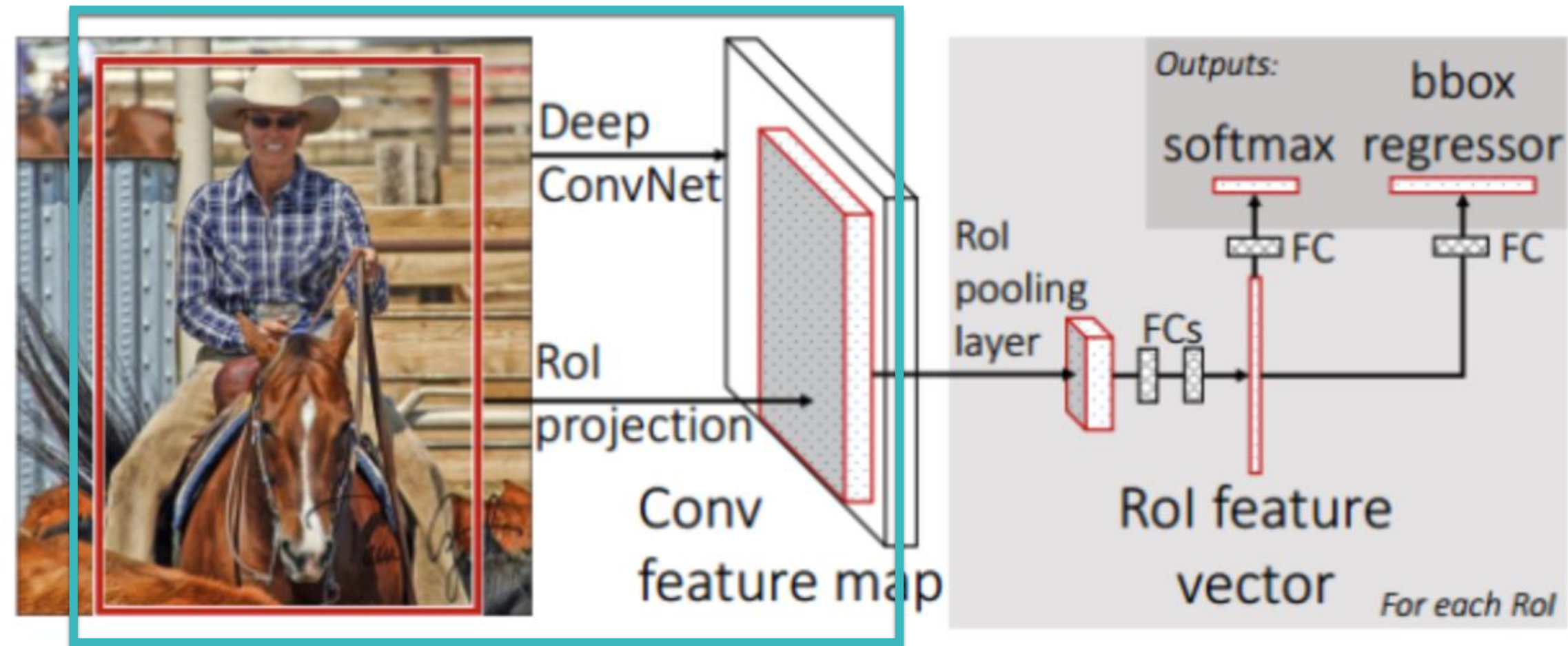
- Extract image regions
- 1 CNN per region (2000 CNNs)
- Classify region-based features
- Complexity: $\sim 224 \times 224 \times 2000$

SPP-net & Fast R-CNN (the same forward pipeline)

- 1 CNN on the entire image
- Extract features from feature map regions
- Classify region-based features
- Complexity: $\sim 600 \times 1000 \times 1$
- $\sim 160\times$ faster than R-CNN

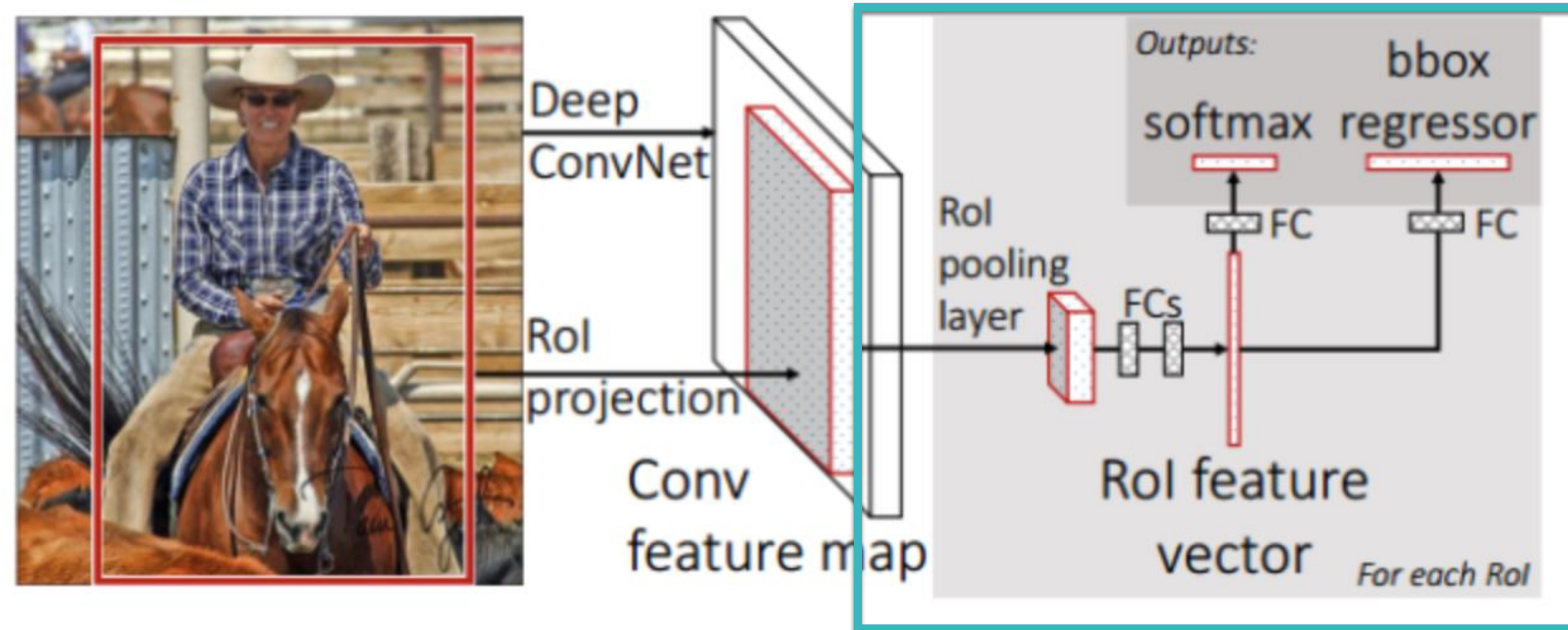
R-CNN에서는 2,000개의 Region Proposal에 대해 계산 많은 시간이 소요됨 , 하지만 Fast R-CNN에서는 ?

Fast R-CNN



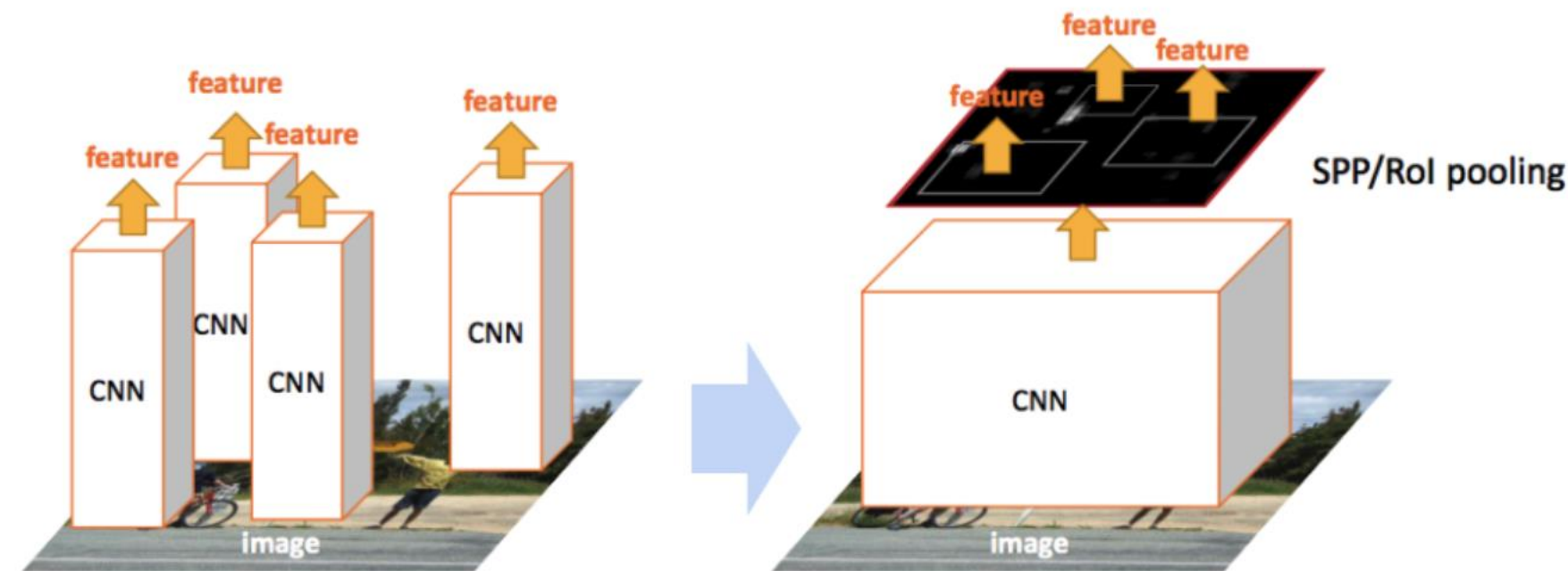
입력 이미지에 한번만 CNN을 적용하고
RoI 풀링으로 객체 판별을 위한 특징을 추출

Fast R-CNN



RoI 영역에 해당되는 부분을 Pooling, 7*7 Feature Map생성
이를 FC 레이어를 거쳐 Object Class와 좌표를 얻어내는 구조

R-CNN vs Fast R-CNN



R-CNN

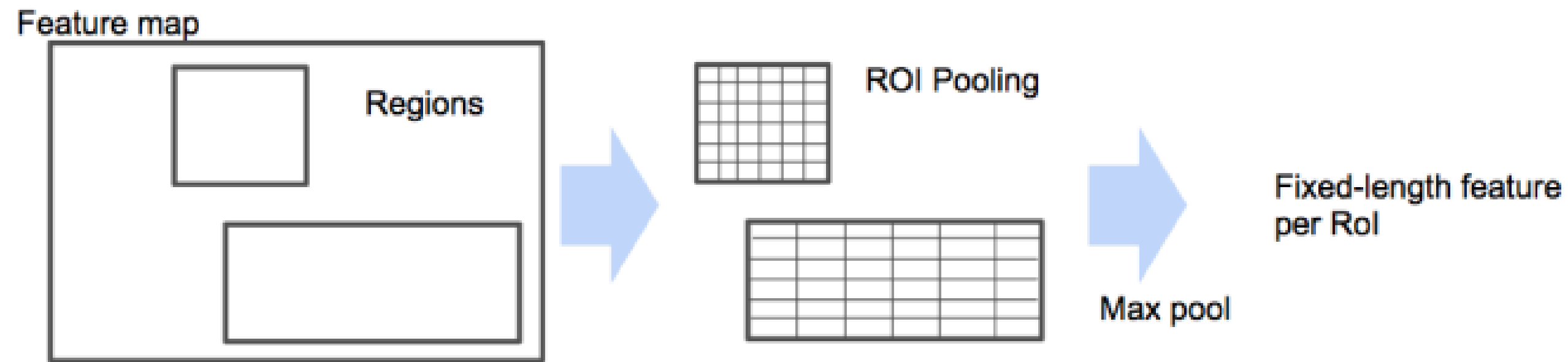
- Extract image regions
- 1 CNN per region (2000 CNNs)
- Classify region-based features
- Complexity: $\sim 224 \times 224 \times 2000$

SPP-net & Fast R-CNN (the same forward pipeline)

- 1 CNN on the entire image
- Extract features from feature map regions
- Classify region-based features
- Complexity: $\sim 600 \times 1000 \times 1$
- $\sim 160\times$ faster than R-CNN

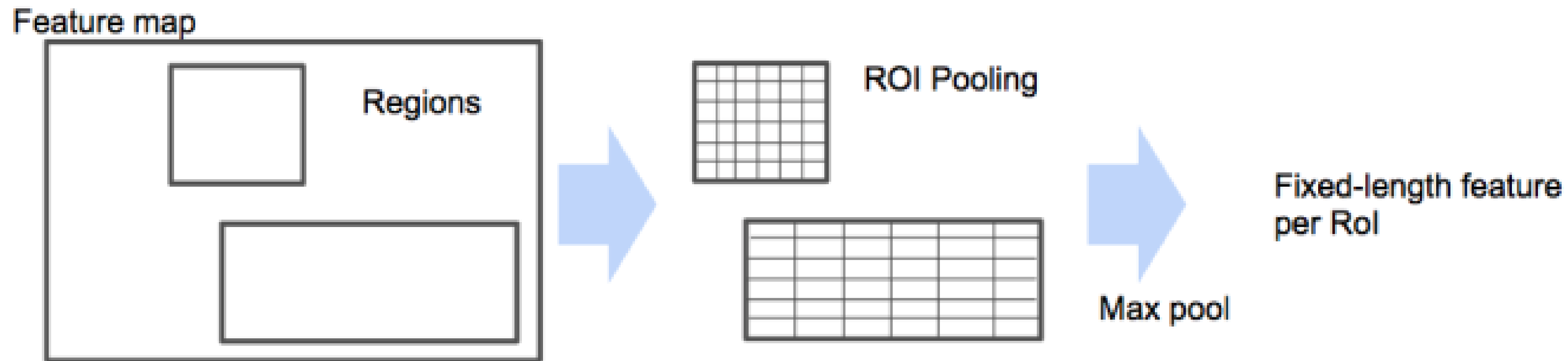
R-CNN: 2,000개의 Region마다 이미지를 Cropping, CNN 연산
Fast F-CNN: 한번만 연산 (Region of Interesting, ROI)
하지만, Wrapping은 사이즈/비율을 조절하기 때문에 **정보의 손실**

Roi를 찾기 위한 Roi Pooling



어떤 Region이 들어오더라도 Max Pooling을 이용하여
동일한 Output feature map의 크기를 갖도록 하는 방법
예) 21×14 Region 이라면, 3×3 Max Pooling (3,2 Stride) 적용

Roi를 찾기 위한 Roi Pooling



어떤 Region이 들어오더라도 Max Pooling을 이용하여
동일한 Output feature map의 크기를 갖도록 하는 방법
예) 21×14 Region 이라면, 3×3 Max Pooling (3,2 Stride) 적용

다음 로드뷰에서의 Fast R-CNN

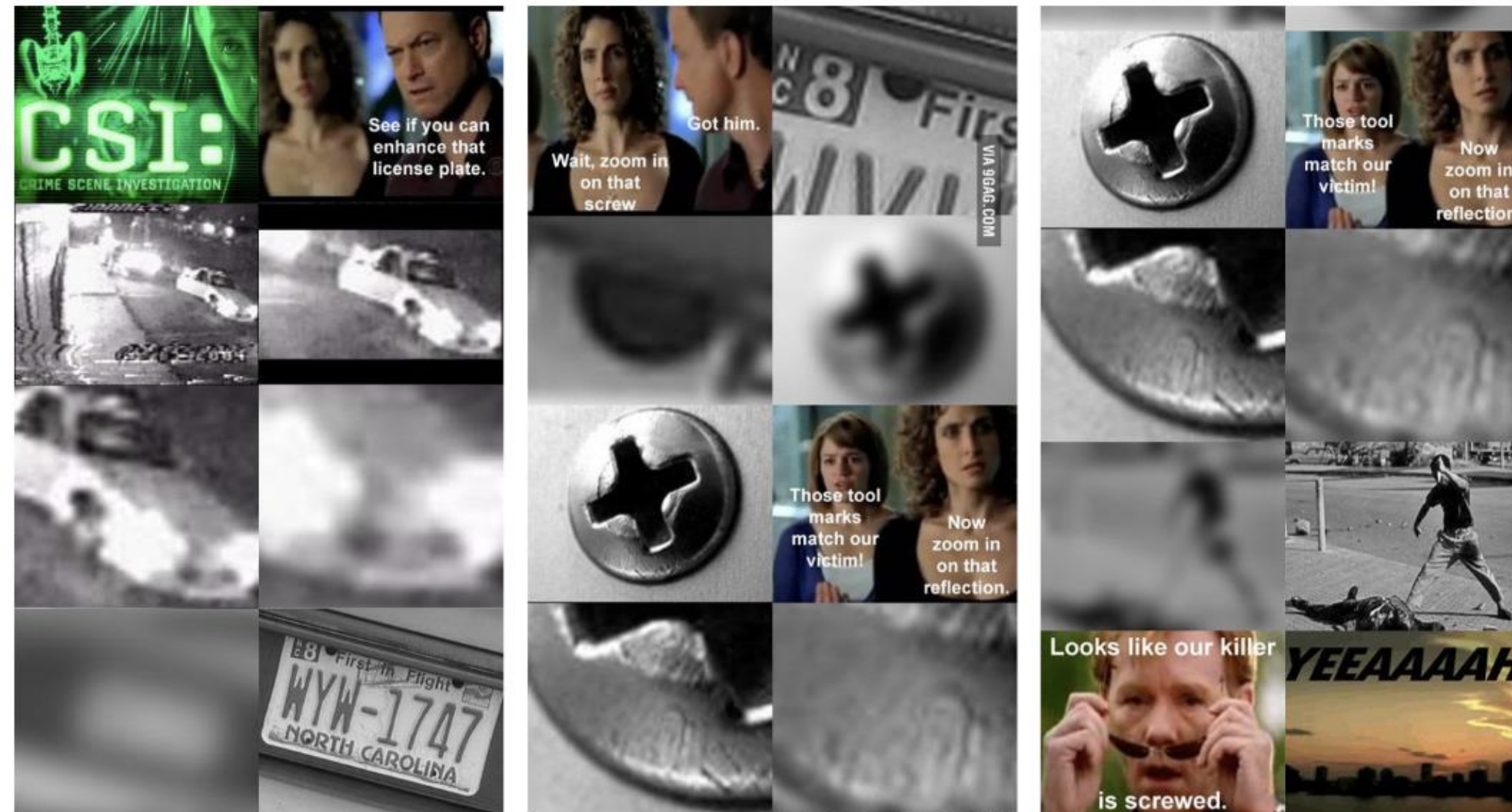


로드뷰 서비스에서는 개인정보의 보호가 필요
차량의 번호를 검출하기 위해 Fast R-CNN 적용

Super Resolution CNN

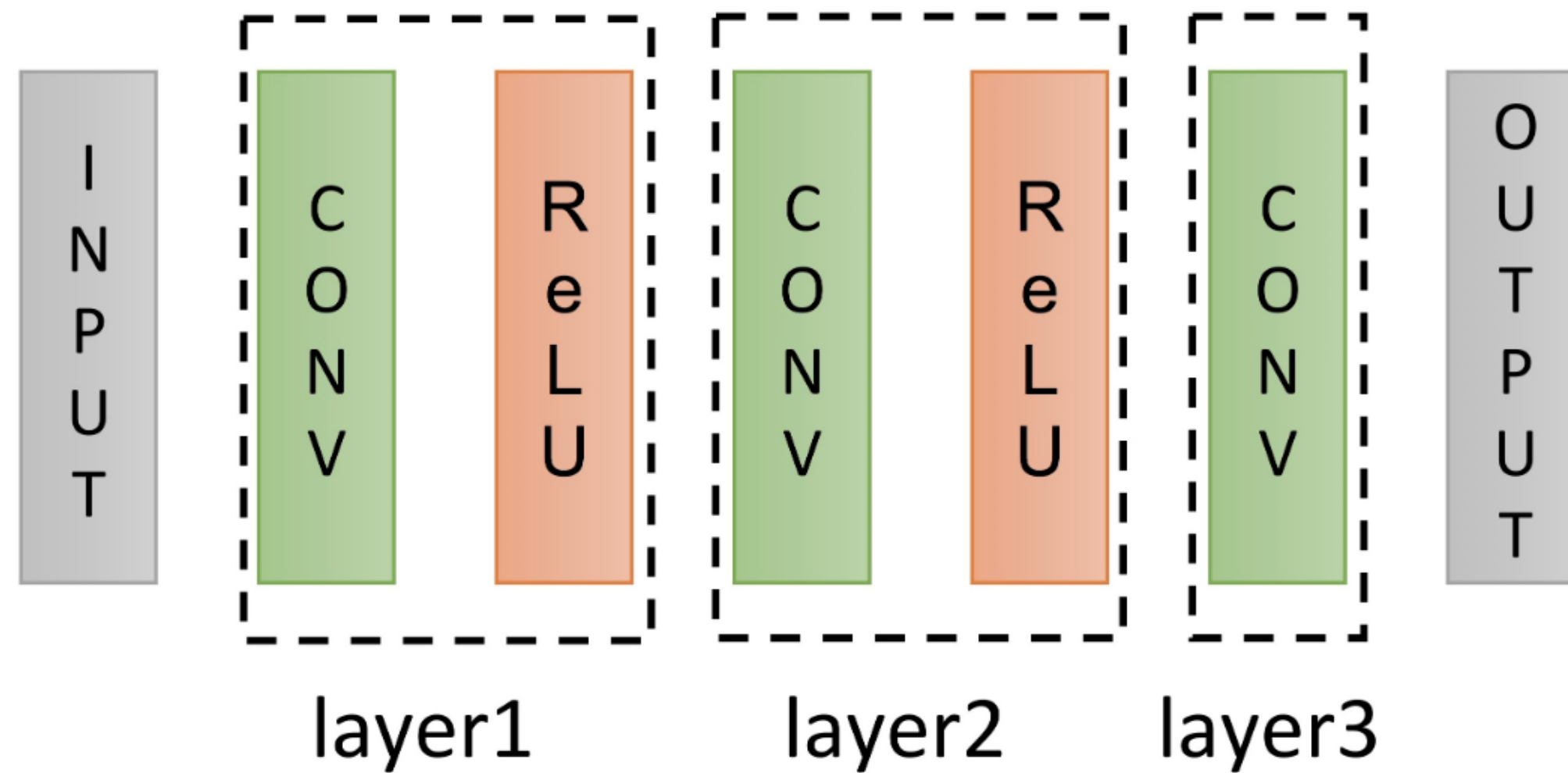
이미지의 품질을 높이는 방법

Super Resolution: 이미지 복원



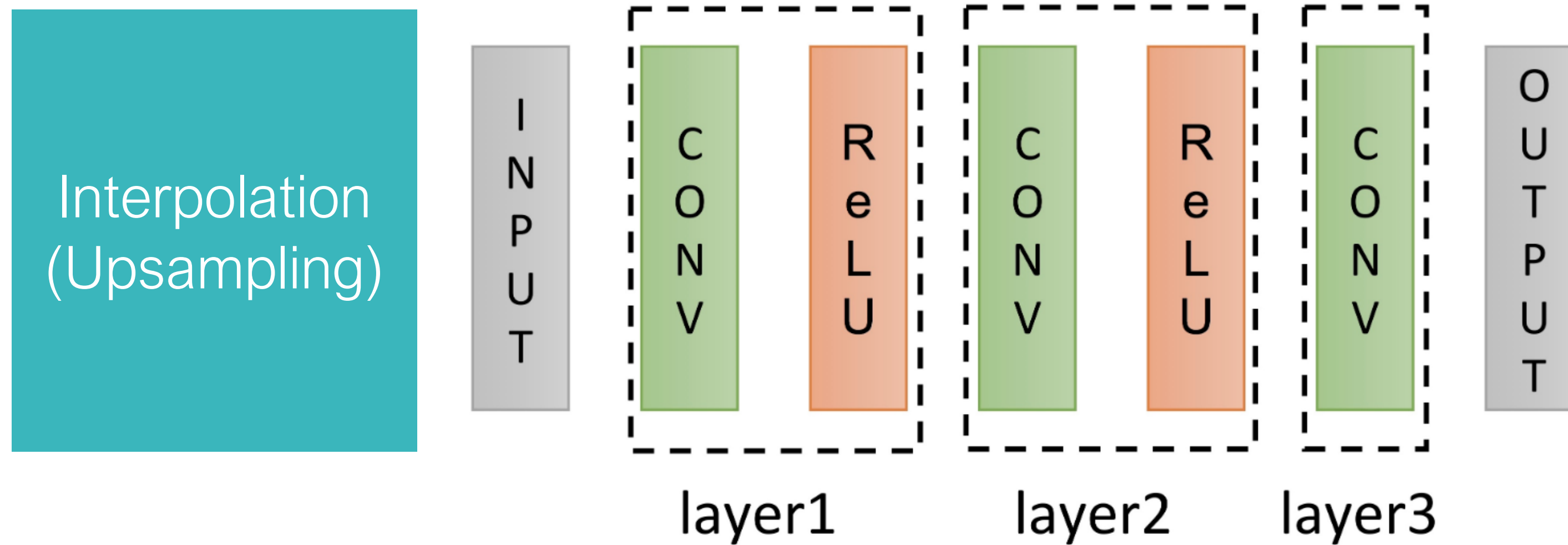
SR은 이미지나 비디오의 해상도를 좋게 하는 영상처리 기술

Super Resolution CNN 구조



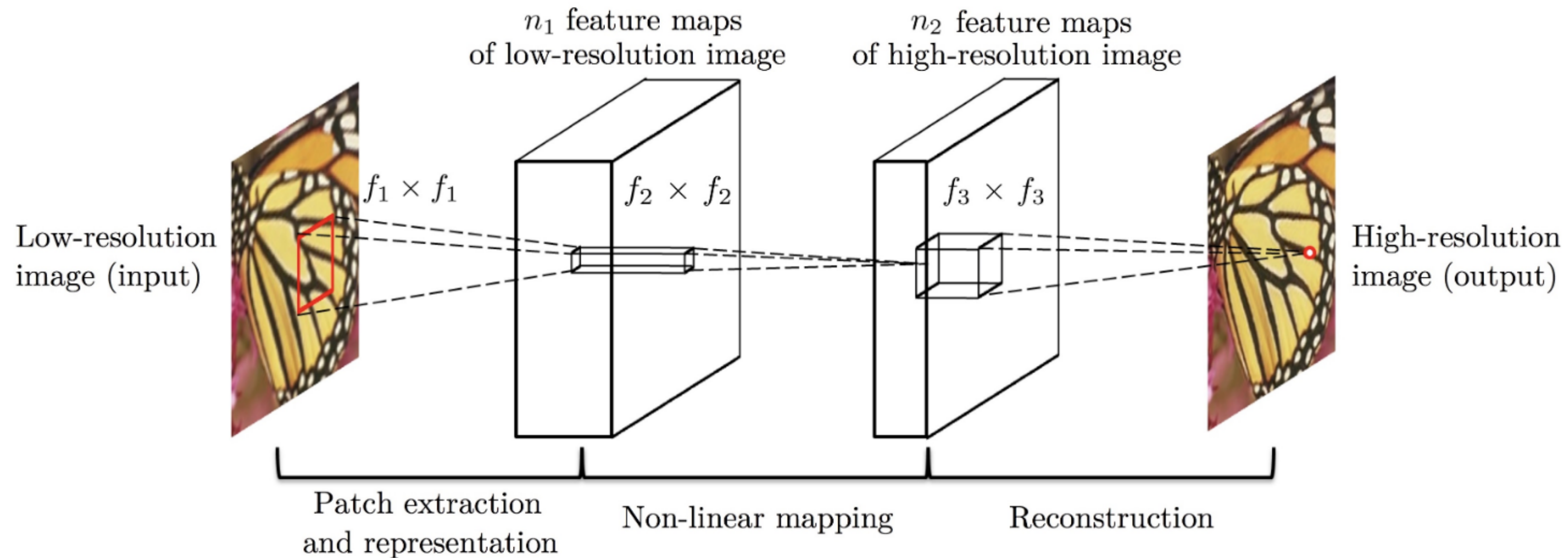
간단한 CNN 구조를 사용하여 문제 해결
학습에는 Mean Square Error (MSE) 사용

SRCNN 동작 순서



CNN을 사용하여 이미지를 연산하기 전에 이미지의 크기를 키움
Bicubic Upsampling 이라는 방법이 사용됨

SRCNN

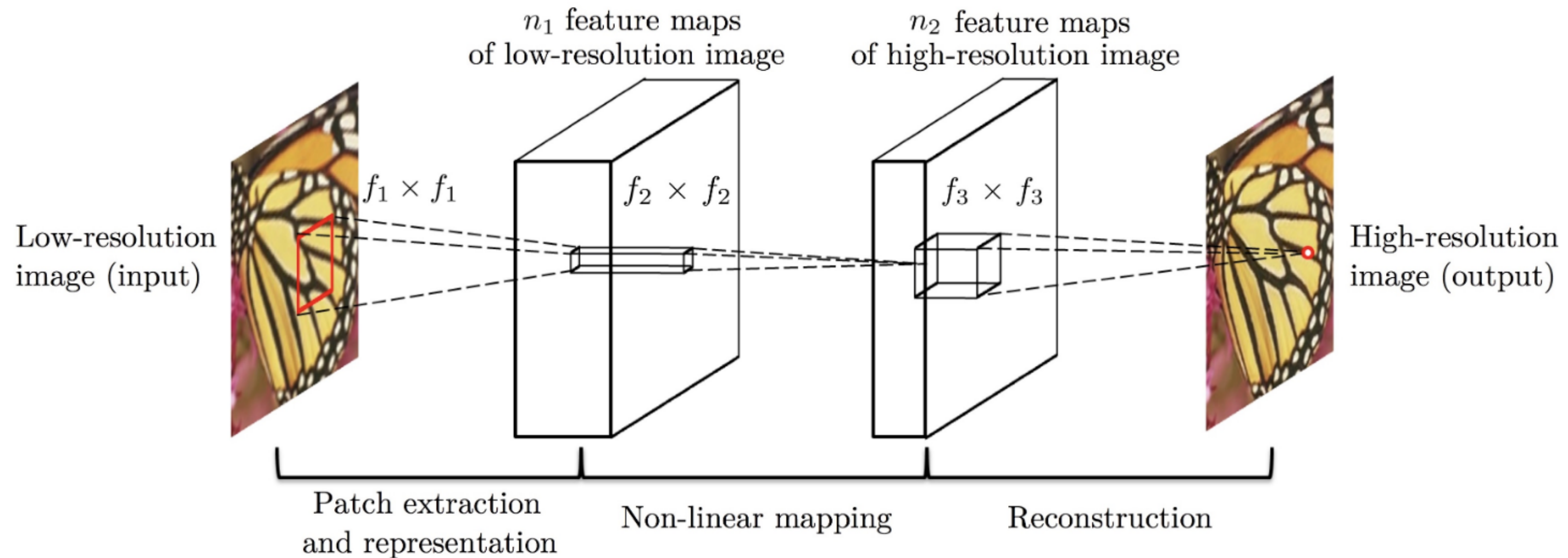


Patch~representation: 저해상도 이미지 YY로부터 patch 추출

Non-linear mapping: 다차원 patch 벡터를 다른 벡터로 mapping

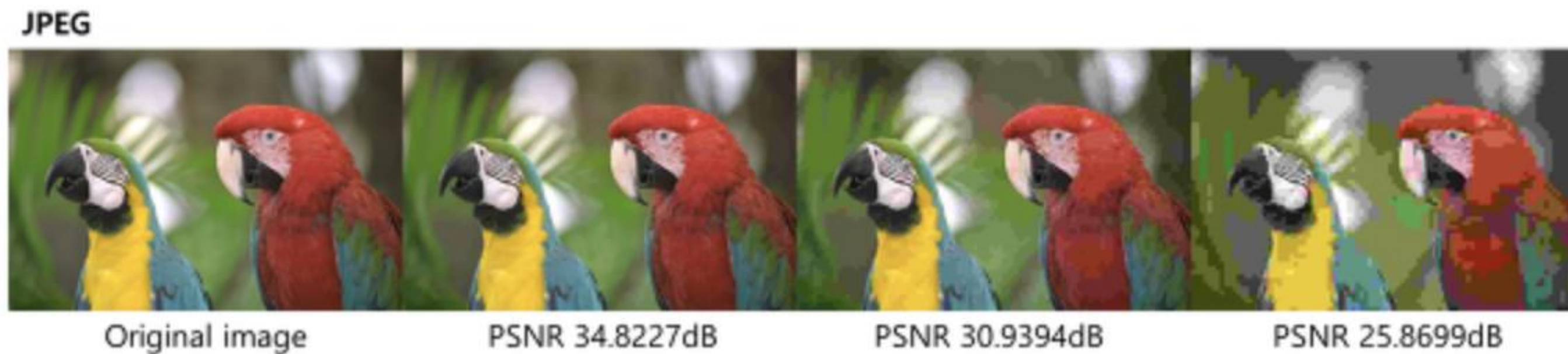
Reconstruction: 다차원 patch 벡터에서 최종 고해상도 이미지 생성

SRCNN



$$\begin{aligned}\text{Layer 1: } F_1(Y) &= \max(0, W_1 * Y + B_1) \\ \text{Layer 2: } F_2(Y) &= \max(0, W_2 * F_1(Y) + B_2) \\ \text{Layer 3: } F_3(Y) &= W_3 * F_2(Y) + B_3\end{aligned}$$

SRCNN에서 사용하는 성능의 지표

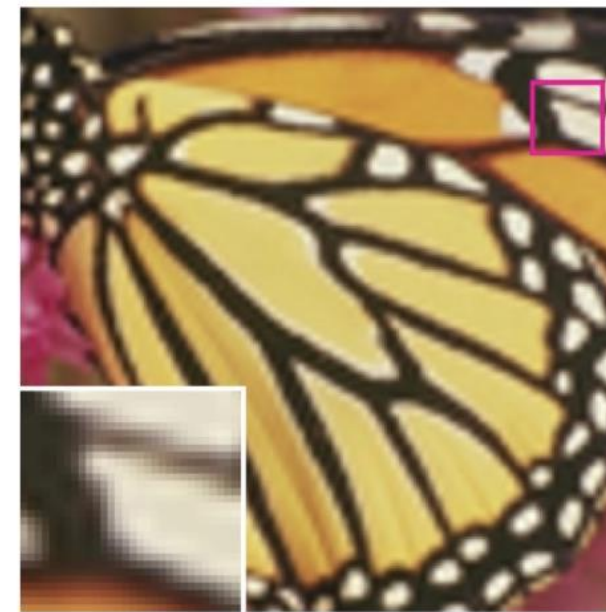


원본 이미지와 비교 대상이 되는 이미지 사이의 수치적 차이
(Peak Signal to Noise Ratio, PSNR)
PSNR의 값 (dB)이 높을수록 좋은 값

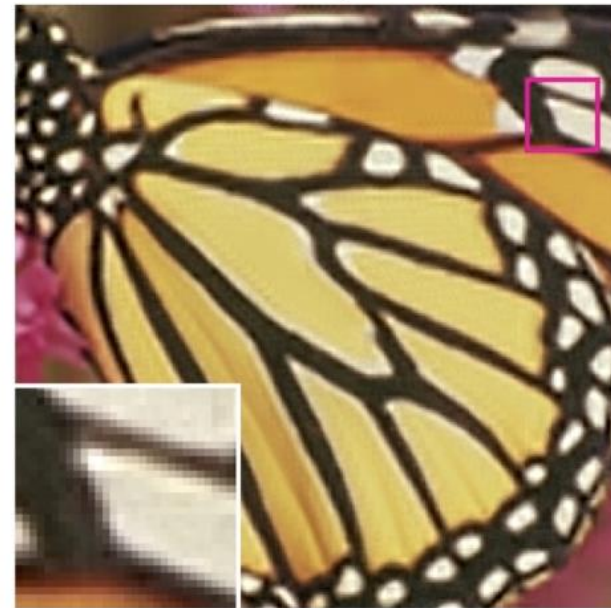
SRCNN 성능 #1



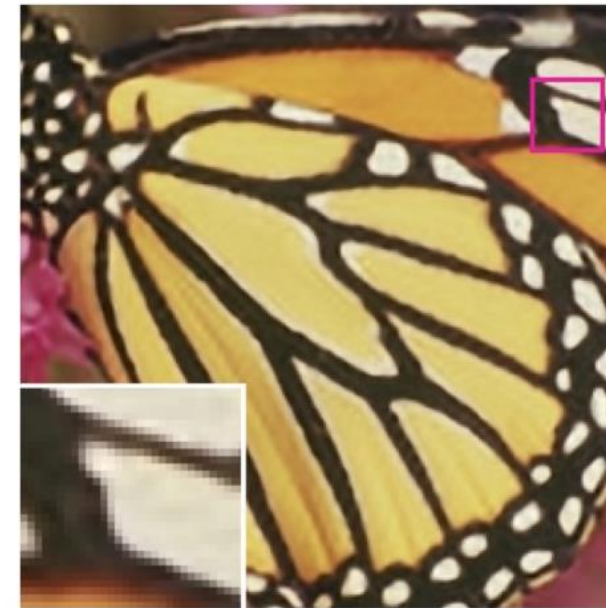
Original / PSNR



Bicubic / 24.04 dB



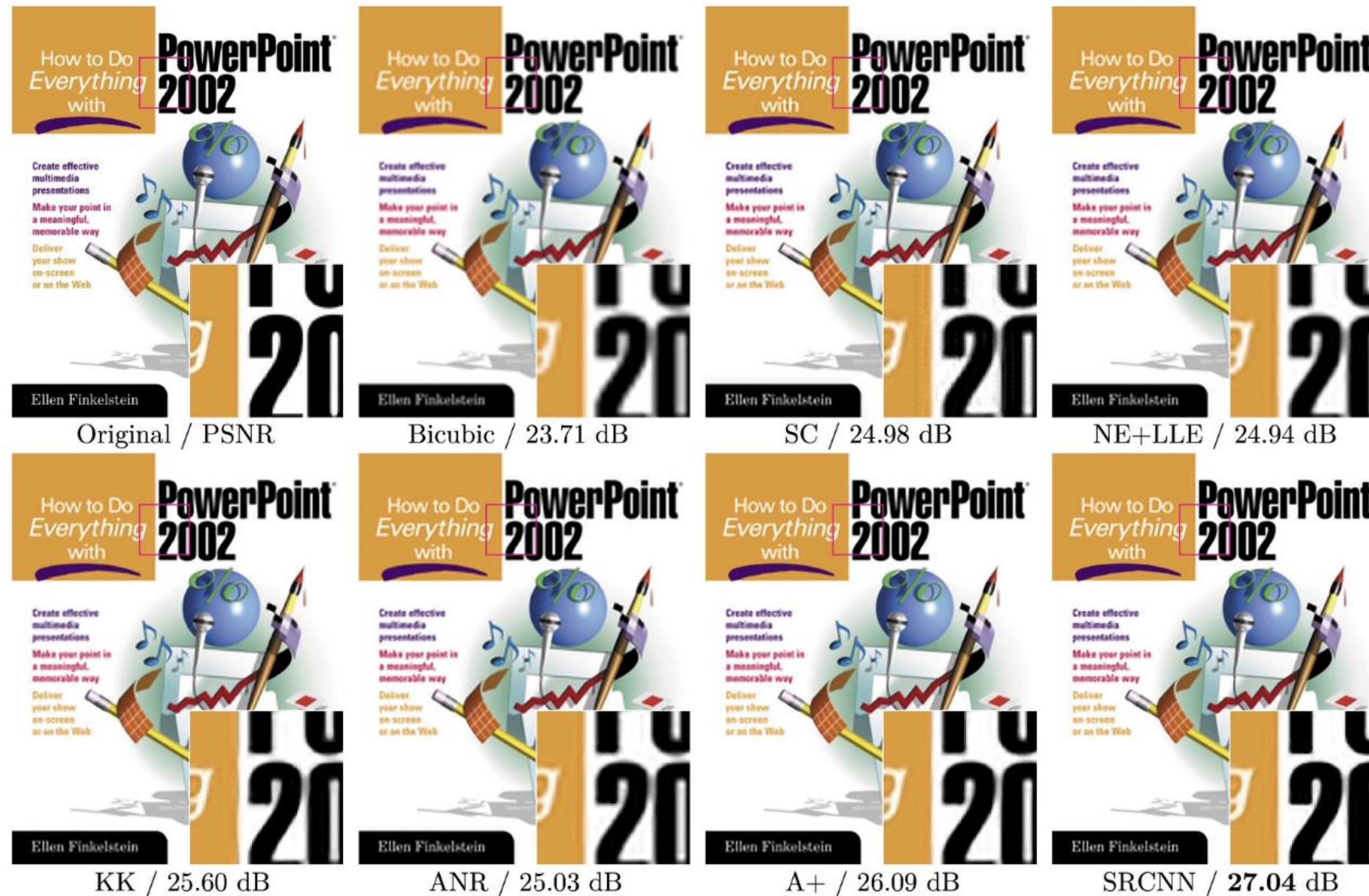
SC / 25.58 dB



SRCNN / 27.95 dB

기존의 연구 방법에 비해 PSNR값이 높음
나비 날개 이미지에서도 SRCNN이 좋음의 확인 가능

SRCNN 성능 #2



SRCNN 성능 #3



Original / PSNR



Bicubic / 26.63 dB



SC / 27.95 dB



NE+LLE / 28.31 dB



KK / 28.85 dB



ANR / 28.43 dB



A+ / 28.98 dB

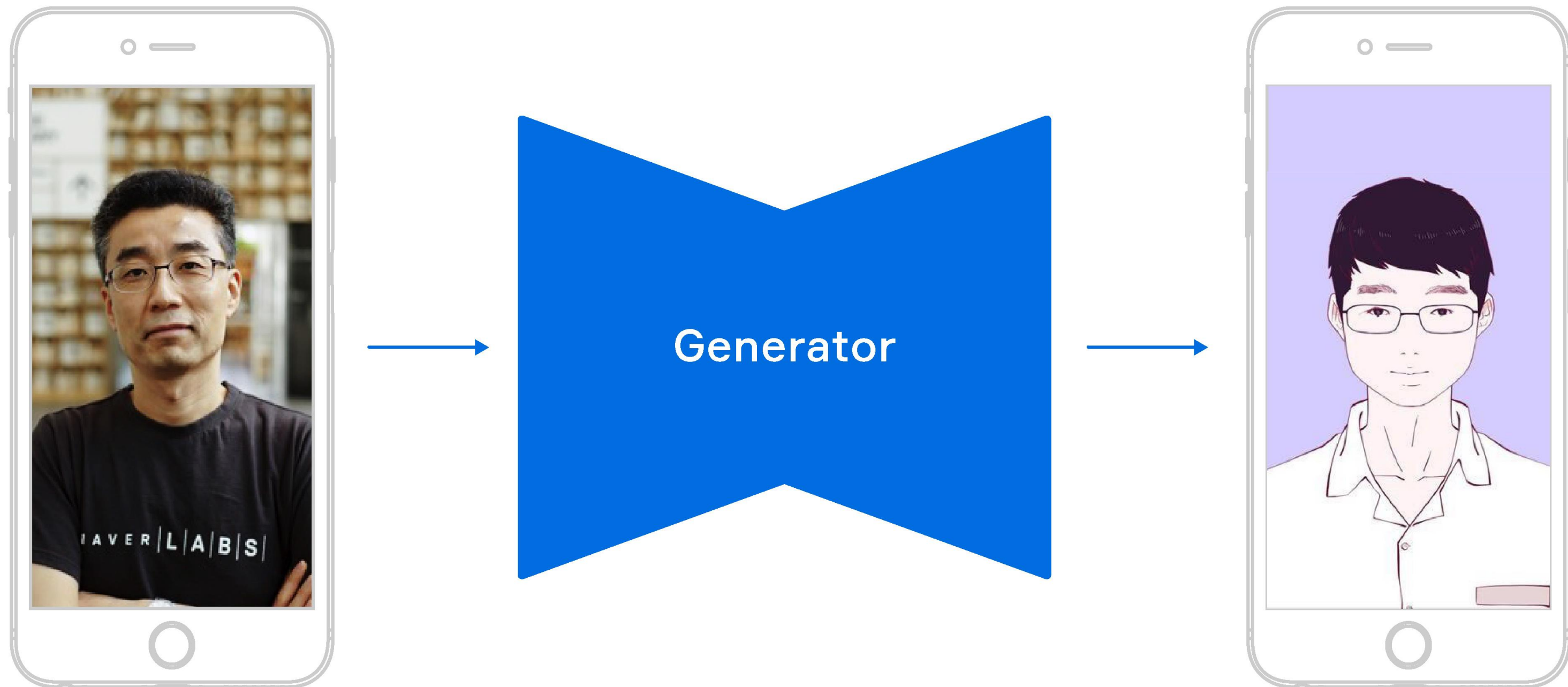


SRCNN / **29.29 dB**

Generative Adversarial Network

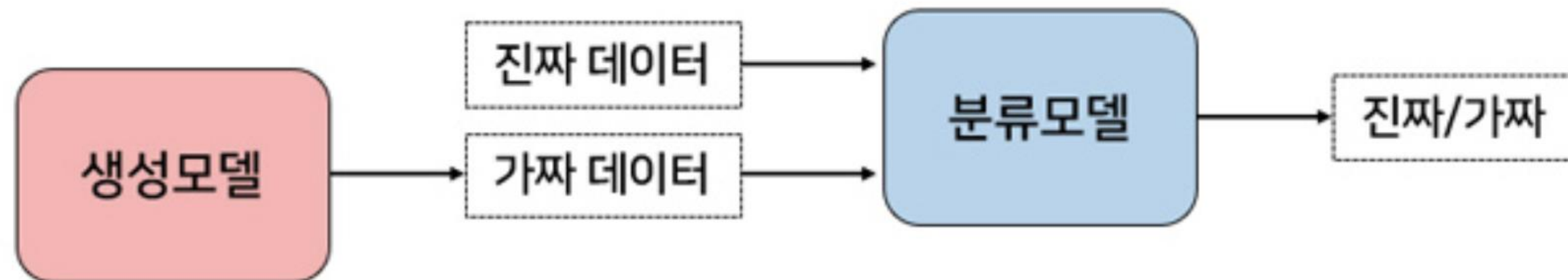
가짜 이미지를 생성하는 신경망

GAN (생성 적대적 신경망)



SR은 이미지나 비디오의 해상도를 좋게 하는 영상처리 기술

GAN 구조



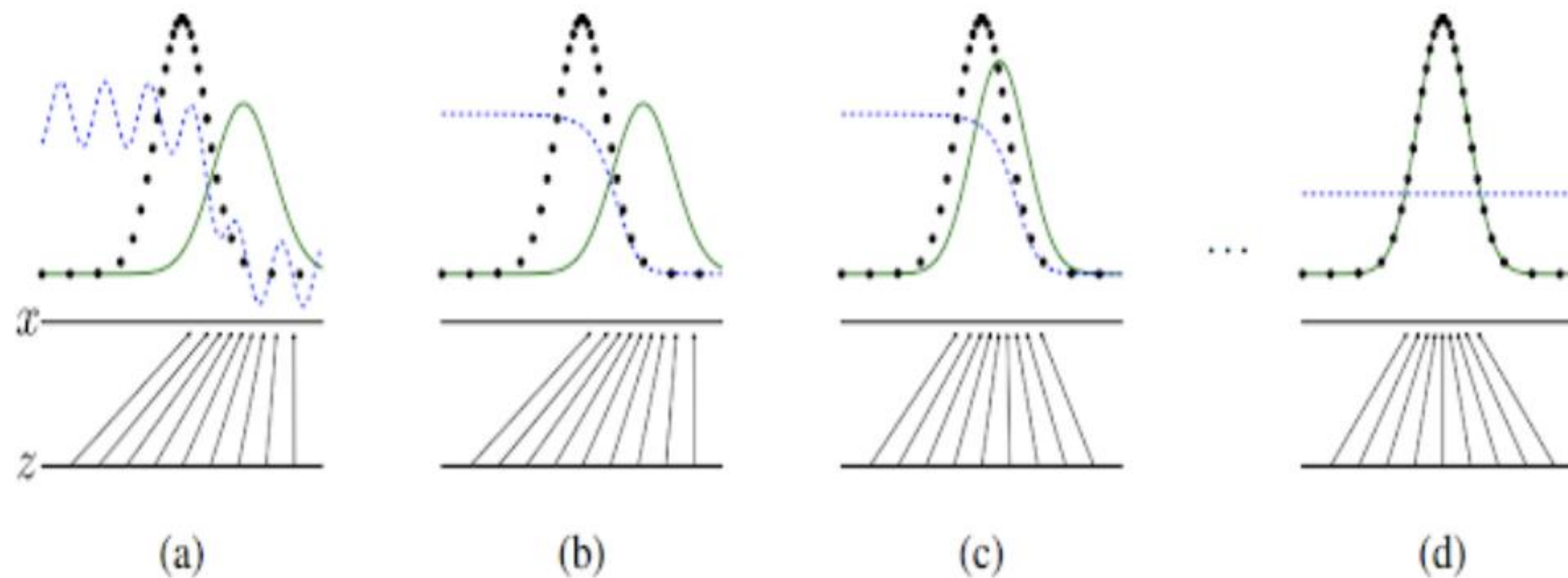
이미지 생성하는 모델, 분류하는 모델이 서로 경쟁하는 신경망
생성모델을 만드는 신경망 (Generator)
분류를 하는 신경망 (Discriminator)

분류모델 vs 생성모델



위조지폐범은 위조 화폐를 만들어 경찰을 속이기 위해 노력
경찰은 진짜 화폐와 가짜 화폐를 완벽하게 구분하려고 함
경찰이 생성 모델이 만든 이미지를 진짜라고 생각하게 하는 것

GAN 학습 방법



검은 점선은 Data Generation Distribution
파란 색은 Discriminator Distribution
녹색 선은 Generative Distribution
Discriminator가 구분하지 못할 때 까지 학습 (50%)

GAN으로 생성한 이미지



1. GAN을 사용하여 생성된 이미지와 실제 이미지를 비교하여 생성된 이미지의 품질을 평가한다.

GAN으로 생성한 이미지



GAN에게 특정한 문장 스타일이나 화법을 학습시켜
이미지를 만드는 것도 가능함

`/* elice */`

문의 및 연락처

academy.elice.io

contact@elice.io

facebook.com/elice.io

medium.com/elice