

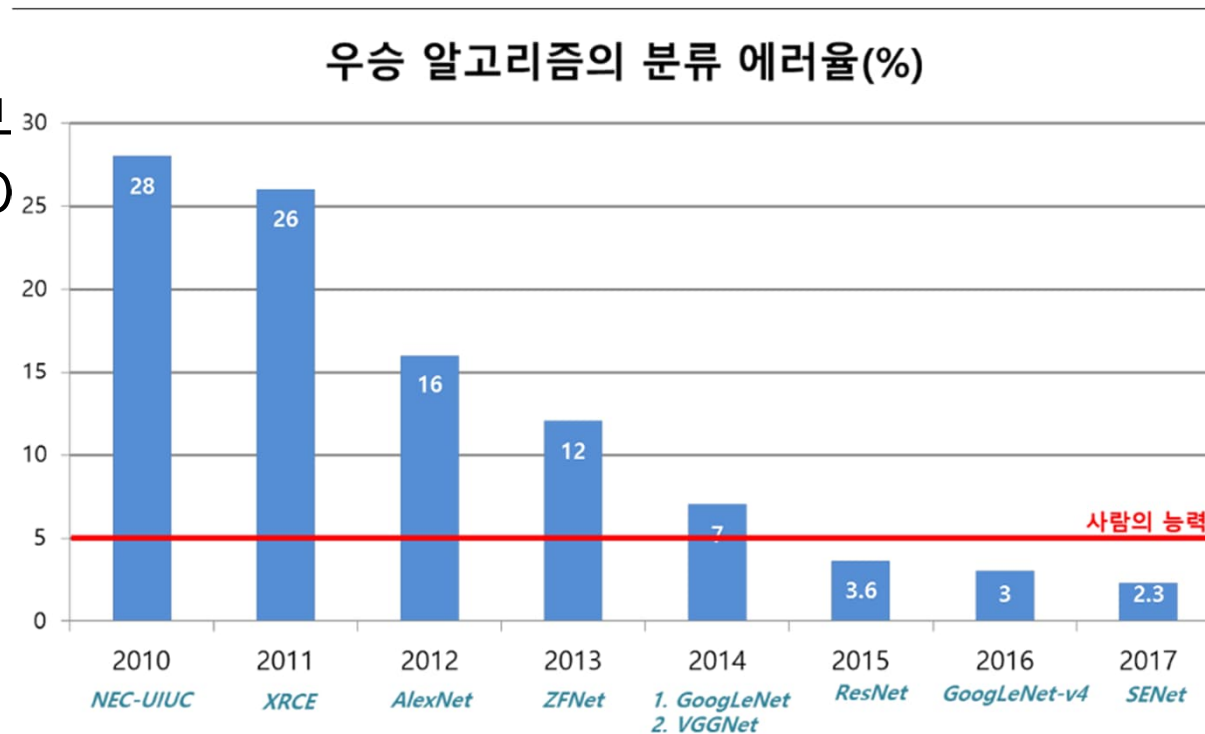
CNN 심화

곽동혁

ResNet

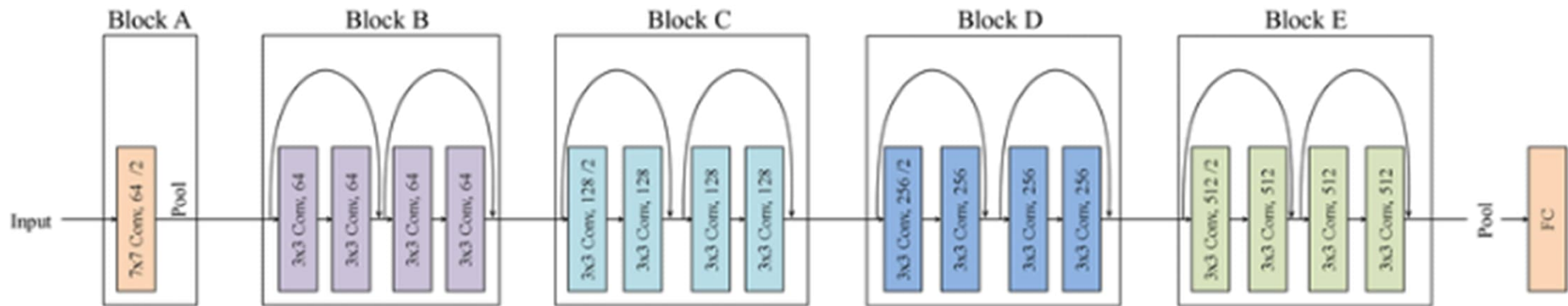
- 2015년도 ILSVRC(ImageNet Large Scale Visual Recognition Challenge)에서 우승한 CNN network

- VGGNet 관련
하였으나 너무
의 영향이 작
발견함



라고
weight
제를

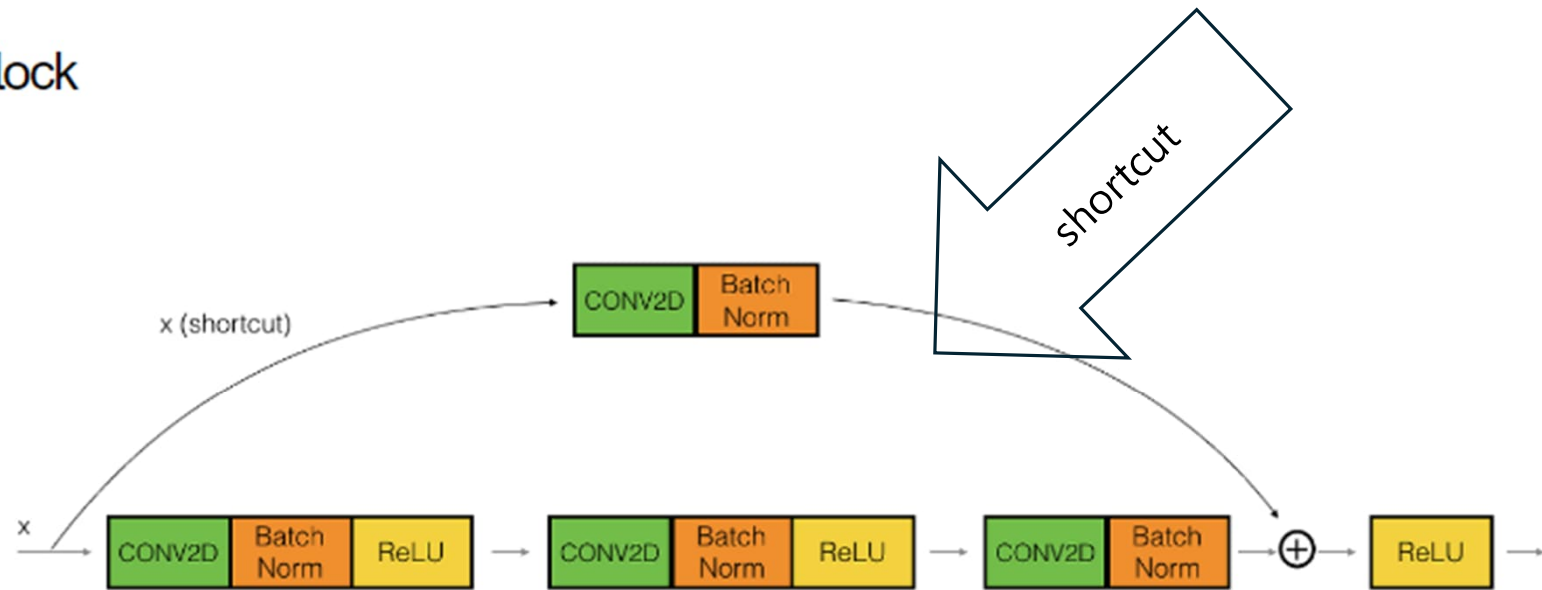
ResNet18 구조



ResNet18 구조

Convolution Block이란

Convolution Block



Convolution Block

Batch normalization

- Internal Covariance Shift
- 단순히 정규분포를 그려고 한다.
- 이를 해결하기 위해 감키는 과정이 batch normalization

Input: Values of x over a mini-batch: $\mathcal{B} = \{x_1 \dots x_m\}$;

Parameters to be learned: γ, β

Output: $\{y_i = \text{BN}_{\gamma, \beta}(x_i)\}$

$$\mu_{\mathcal{B}} \leftarrow \frac{1}{m} \sum_{i=1}^m x_i \quad // \text{ mini-batch mean}$$

$$\sigma_{\mathcal{B}}^2 \leftarrow \frac{1}{m} \sum_{i=1}^m (x_i - \mu_{\mathcal{B}})^2 \quad // \text{ mini-batch variance}$$

$$\hat{x}_i \leftarrow \frac{x_i - \mu_{\mathcal{B}}}{\sqrt{\sigma_{\mathcal{B}}^2 + \epsilon}} \quad // \text{ normalize}$$

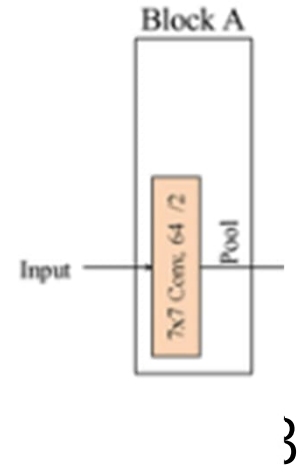
$$y_i \leftarrow \gamma \hat{x}_i + \beta \equiv \text{BN}_{\gamma, \beta}(x_i) \quad // \text{ scale and shift}$$

Algorithm 1: Batch Normalizing Transform, applied to activation x over a mini-batch.

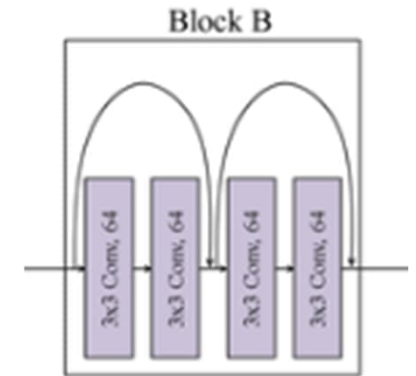
배치 정규화의 입력 및 출력 값

ResNet18 구조

- Input : 224x224x3(RGB) size
- Block A : 7x7 size kernel 64개를 이용하여 입력이 max pooling 한다. 이때 stride = 2, zero padding = 1 하면 output의 size는 112 x 112 x 64가 된다.(zero padding이라는 것은 가로 세로의 크기가 6씩 커짐)
- (참고로 각각의 kernel은 R,G,B 각각에 적용되어 곱값을 모두 더한다.)
- 이후 max pooling을 stride = 2 로 진행하여 56x56x64로 output을 만든다.

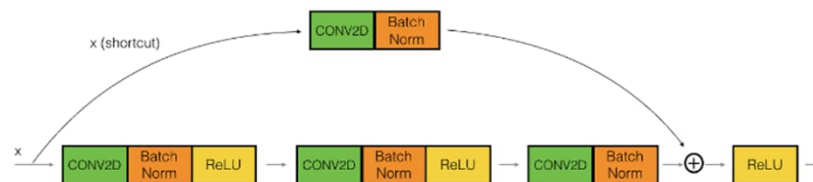


ResNet18 구조



- Input : 56x56x64size
- Block B : 3x3 size kernel 64개를 이용하여 입력이미지를 convolution 한다. 이때 stride = 1, zero padding = 1으로 진행하면 output의 size는 56 x 56 x 64가 된다.(zero padding = 1이라는 것은 가로 세로의 크기가 2씩 커짐)
- 아래 사진 과정을 두번 거친다. Output size는 유지된다.

Convolution Block

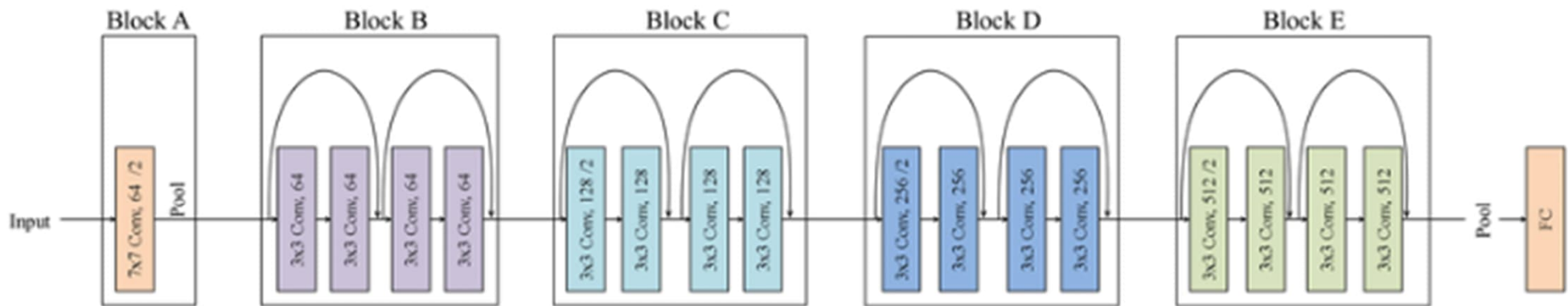


Convolution Block

ResNet18 구조 : 블록 C 이후

$$\sigma(\mathbf{z})_i = \frac{e^{z_i}}{\sum_{j=1}^K e^{z_j}}$$

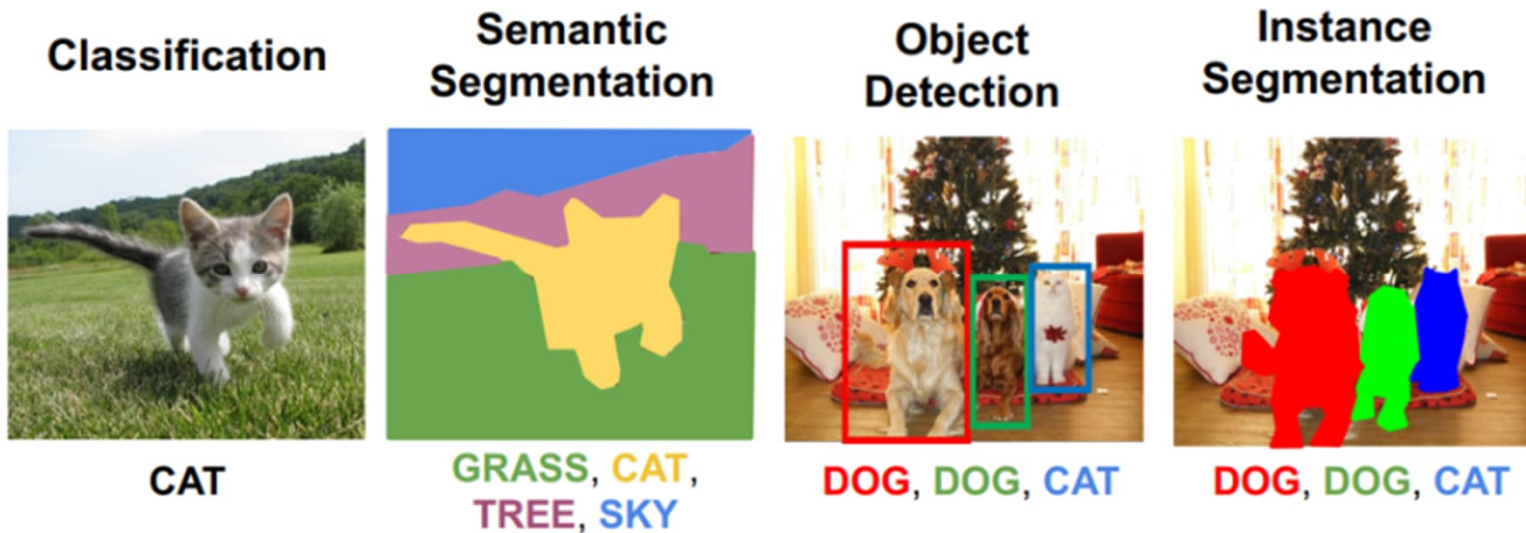
- 비슷하게 C,D,E 진행 → 이후 output 은 7x7x512
- adaptiveAverage pooling 이후 출력층 1000개 뉴런과 연결 후 softmax function을 activation function으로 이용하여 분류



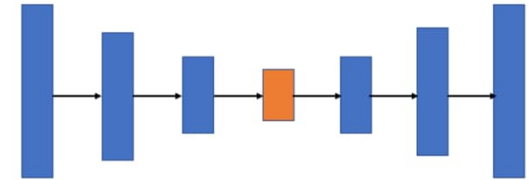
ResNet18 구조

U-Net : 의료용 이미지에 주로 많이 사용

- Image segmentation : 픽셀 단위 분류

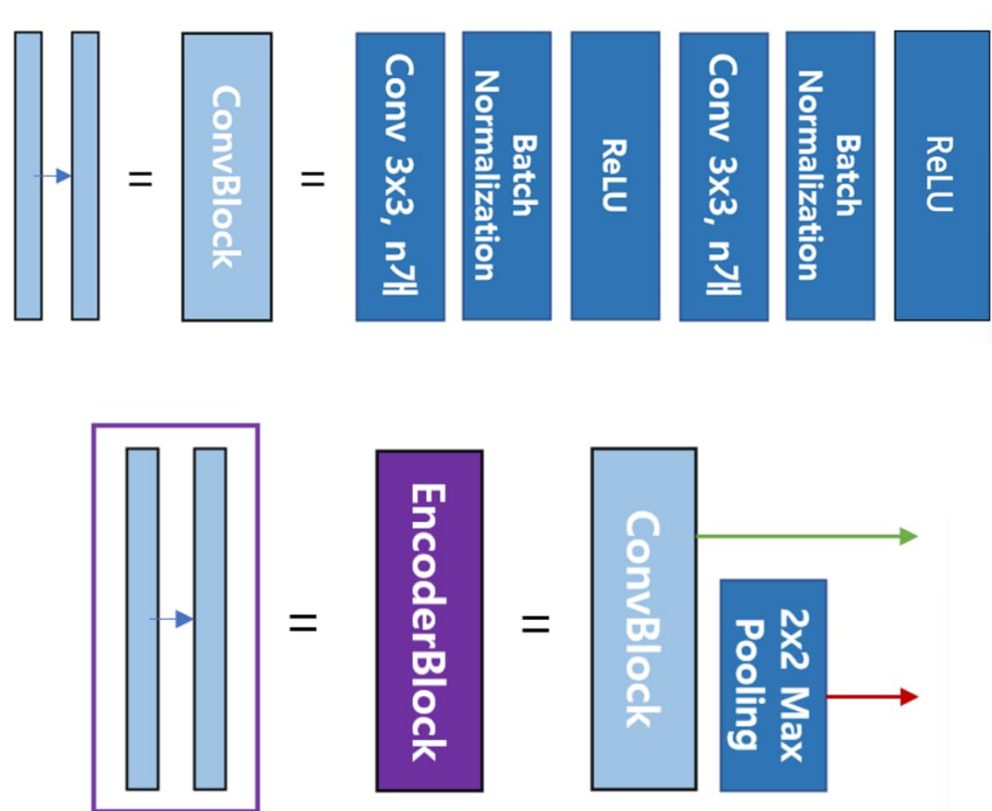
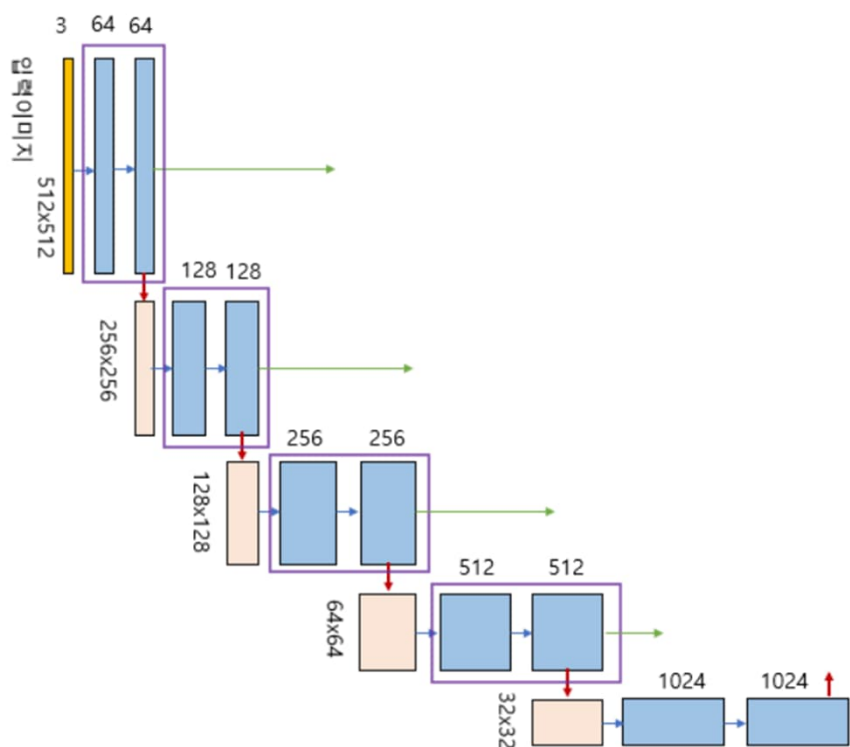


U-Net



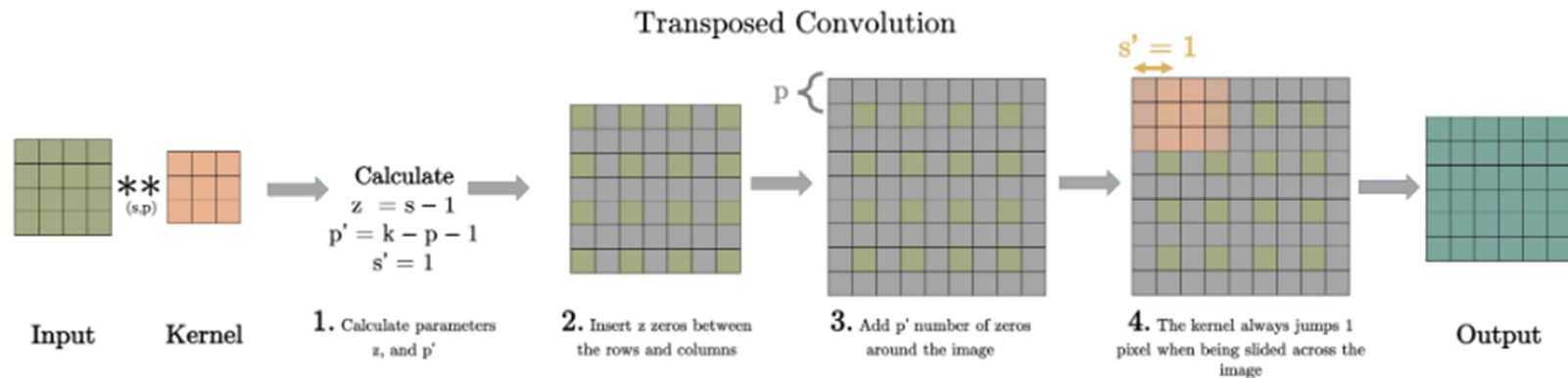
- Encoder-decoder 기반 모델
- 인코딩 : 채널 증가, 차원 축소
- 디코딩 : 채널 축소, 고차원 이미지 복원
- 인코딩 단계와 디코딩 단계를 합치는 방법 사용

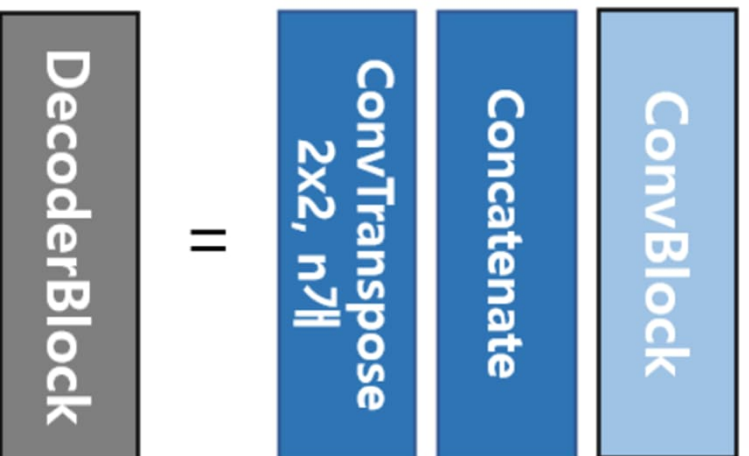
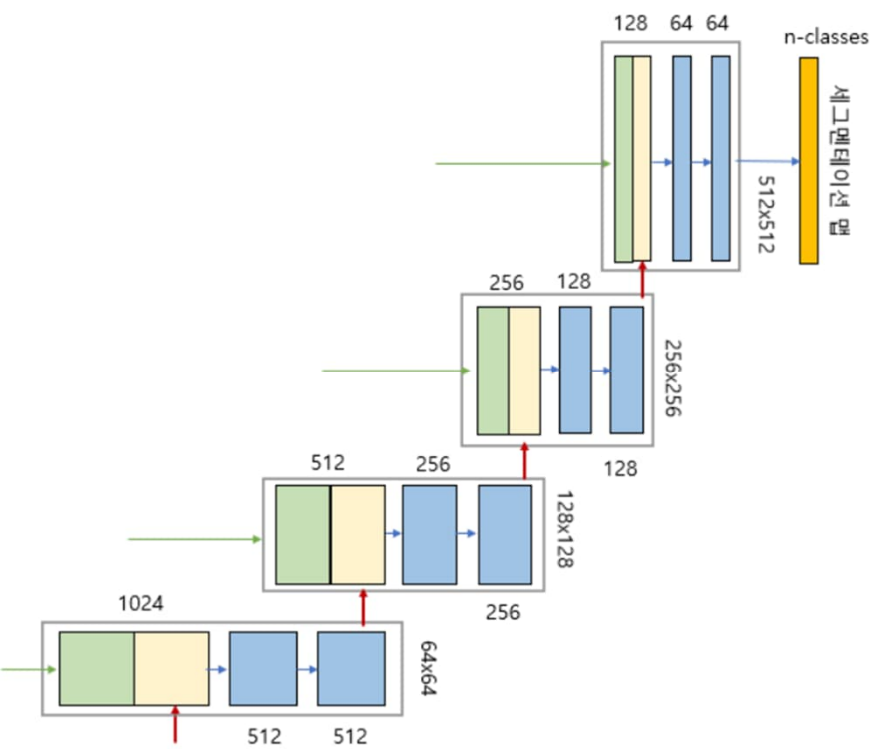
인코딩 단계(contracting path)

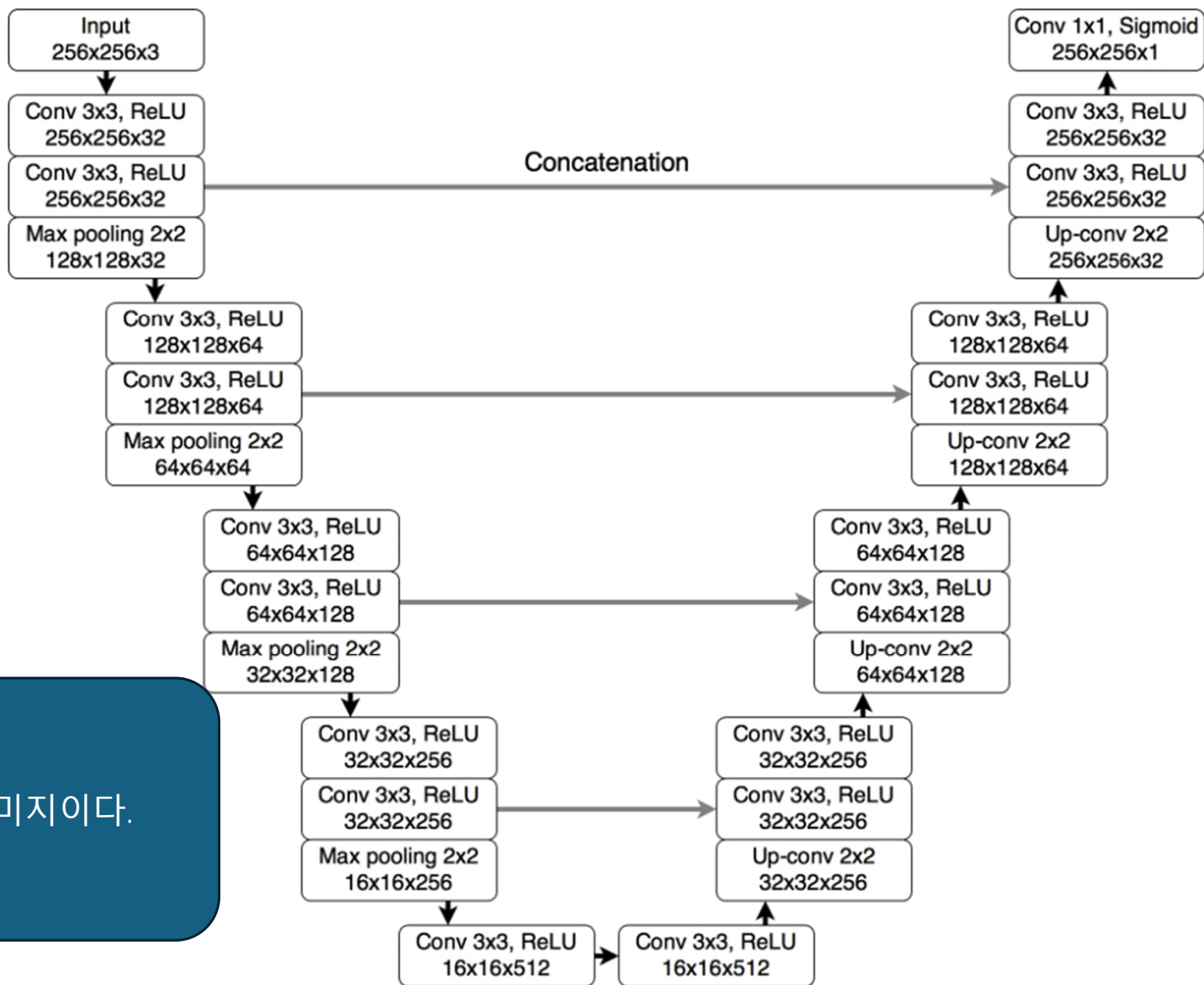


디코딩 단계(Expanding Path)

- Transposed convolutional layer : output의 차원이 input의 차원보다 크다.







결괏값도 이미지이다.