

이미지 분할(Image Segmentation)

출처: <https://wikidocs.net/148872>

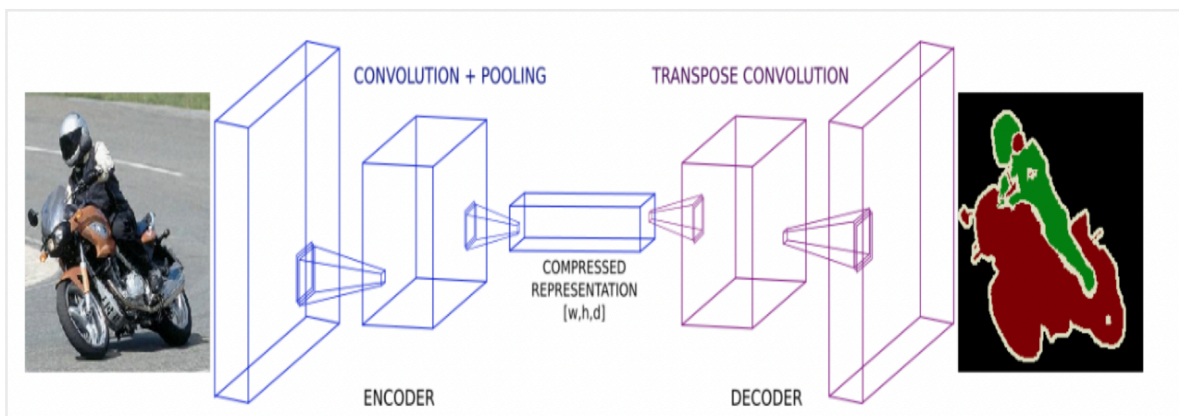
Def: 입력 이미지에서 각 클래스(클래스는 지정 가능)를 구분하는 task

이미지 분할은

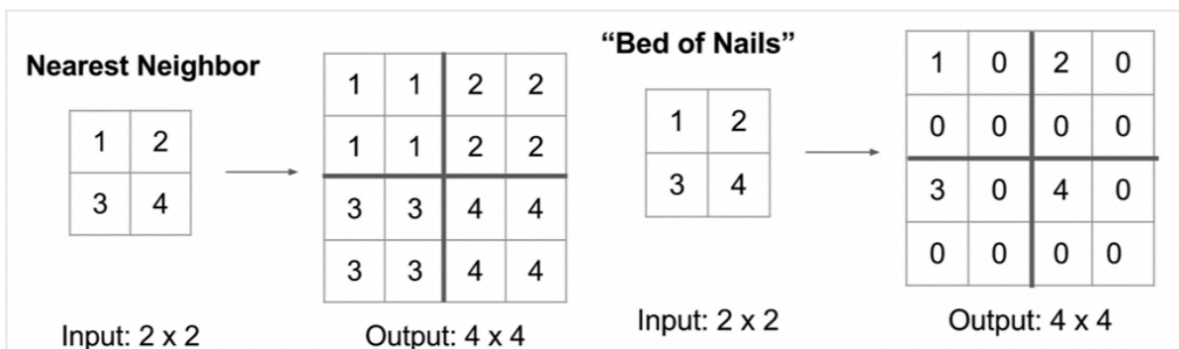
- Semantic segmentation : 각 픽셀이 어떤 클래스에 속하는 지 구분하는 task
- Instance segmentation : 더 나아가 같은 사물 안에서 서로 다른 객체(instance)를 구분하는 task

로 나뉨

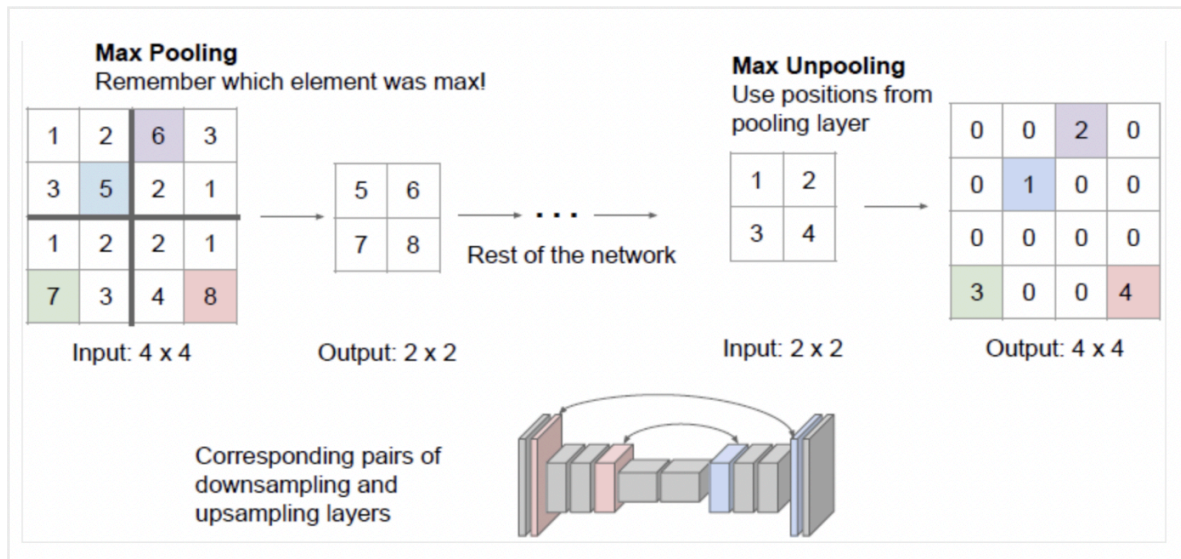
기본구조 : encoder & decoder



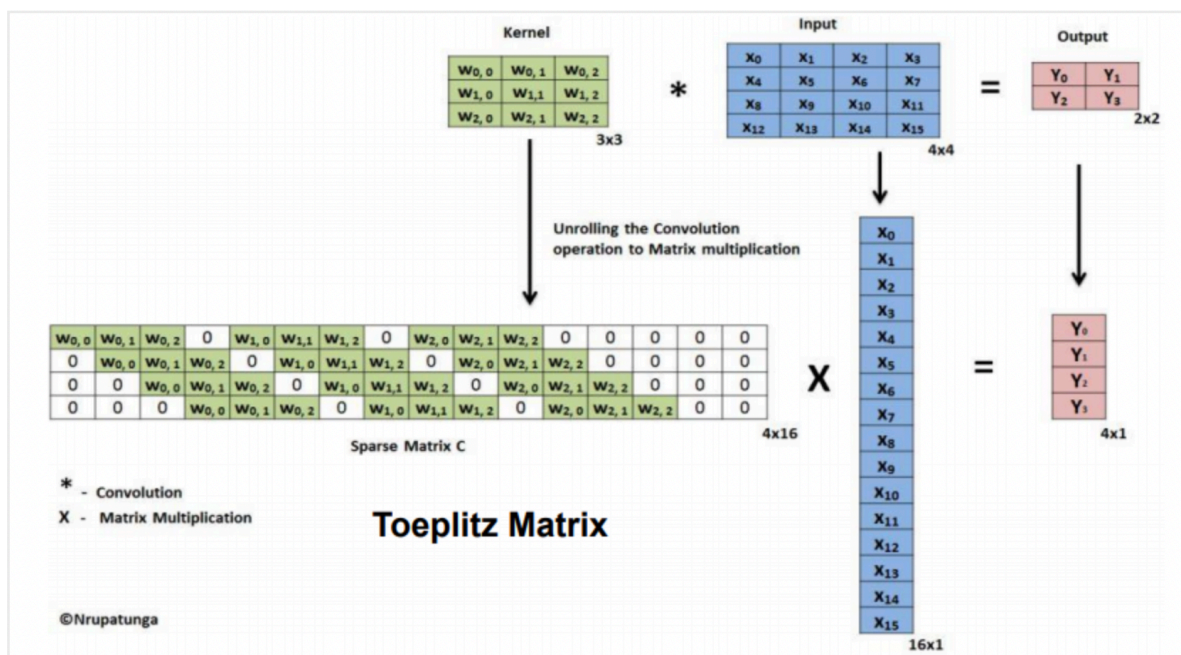
1. 입력 이미지의 크기(너비, 높이)를 줄이고 채널 수를 늘려 피처의 개수 증가시킴
(encoder, down-sampling)
 - pooling(주로 max-pooling, 일반적인 convolution에서는 아웃풋에 채널 없음)
 - dilated (A trous) convolution(구멍이 있는 convolution, dilation rate 지정 가능)
 - depth wise convolution(채널별로 다른 필터 사용, 아웃풋에 채널층 존재)
 - depth wise separable convolution(depth wise convolution의 아웃풋의 채널을 1-d로 합침)
2. 이미지의 크기를 원래 입력이미지 크기로 회복시키고, 채널수는 클래스의 사이로 맞추어 **seg. map** 생성(decoder, up-sampling)
 - unpooling



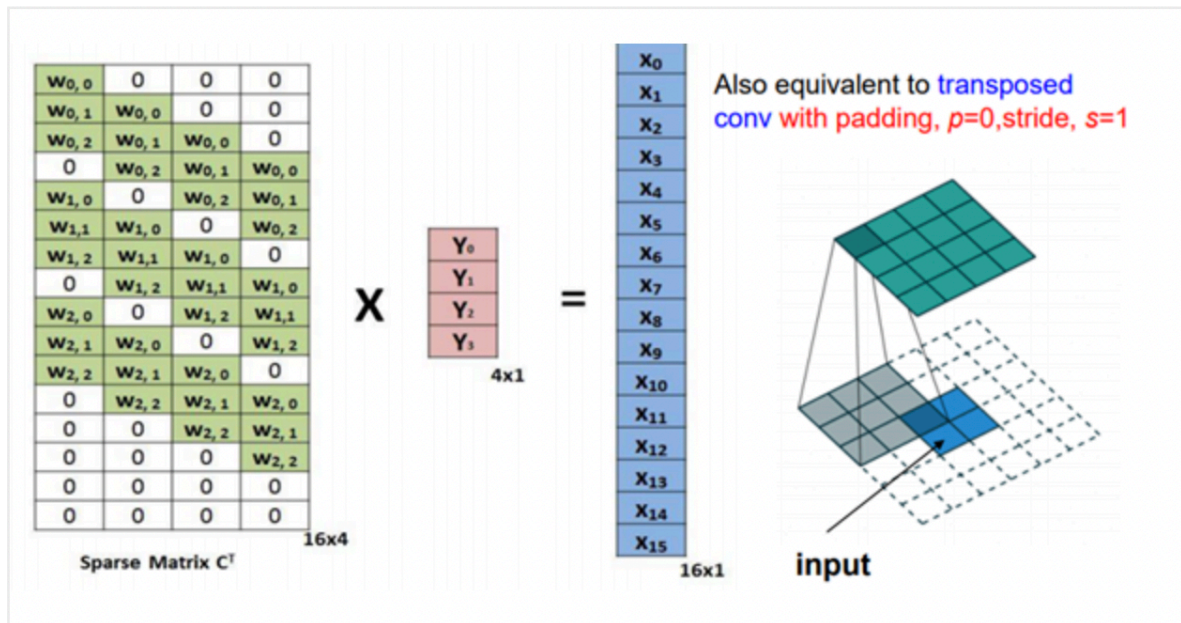
- max unpooling



- bilinear interpolation
- **deconvolution**(filter 행렬의 역행렬 이용)
- **transposed convolution**(backward strided convolution)(학습을 통해(?) 전치하여 새로운 f 찾음)



이게 기존 convolution 방식이라면



이게 transposed convolution 방식

Image segmentation 모델 종류

- FCN (2015 CVPR, 기반이 되는 모델, 전이학습 사용, 기존 분류모델의 FC 층 변형, transpose convolution)
- SegNet(skip architecture를 사용해도 조금 부정확한데 이를 해결하기 위해 사용, FCN보다 느리지만, 성능 좋음, large size class에서 높은 정확도, small size class에서 낮은 정확도)
- U-Net(skip architecture를 확장해서 사요함, 부족한 데이터 양일 때 preprocessing, data augmentation을 통해 문제 극복, 특정 데이터에 한정됨)
<https://www.kaggle.com/code/wangmo/self-driving-cars-road-segmentation-task/notebook>
- DeepLab V3+()