

Semantic Segmentation

- pixel-level classification
- 이미지에 있는 모든 픽셀에 대한 예측을 하는 것

=> **Dense Prediction**

- Semantic Segmentation을 위해 제안된 딥러닝 모델

FCN

DeepLav v1, v2

U-Net

...

- 자율주행에서 활용
- convolutionization

flatten layer 없이 conv 연산으로 label vector 계산

파라미터 개수는 기존 flatten 후 label과 동일

heat map 생성 가능

- **FCN(Fully Convolutional Network)**

최초의 pixelwise end to end 예측모델

convolutional layer만 사용

기존의 CNN 모델 뒤쪽에 있는 FC layer를 1×1 convolutional layer로 대체

- output dimensions reduced by subsampling
- upsampling을 통해 다시 output 사이즈 키워줌
- padding을 적절히 줘서 deconvolution을 사용해 upsample
- **Convolutionalization**

FC layer 사용 X, 1×1 convolutional layer 사용

- FC layer

고정된 입력 크기

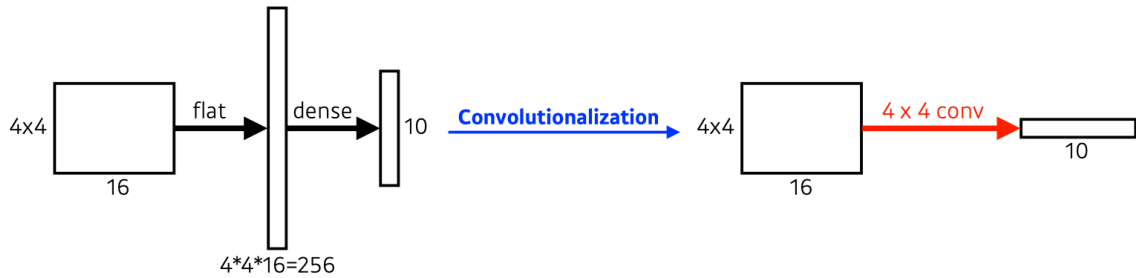
이미지의공간적 정보 소실

- convolutional layer

각 channel의 이미지 위치마다 flattening하여 얻는 벡터들을 fully-connected

=> same as 1×1 convolution

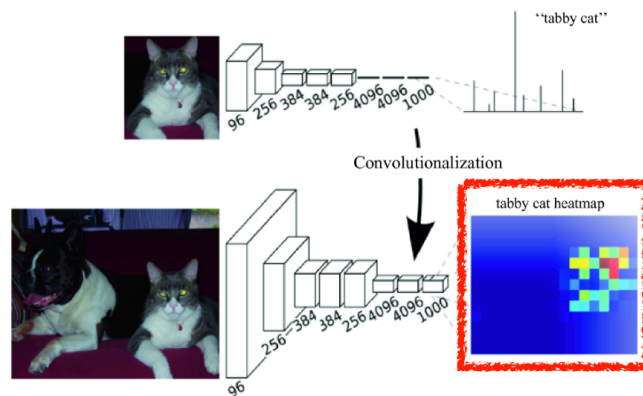
입력사이즈에 Independent => 입력사이즈 제한 X, 공간 정보 보존



● # of parameters

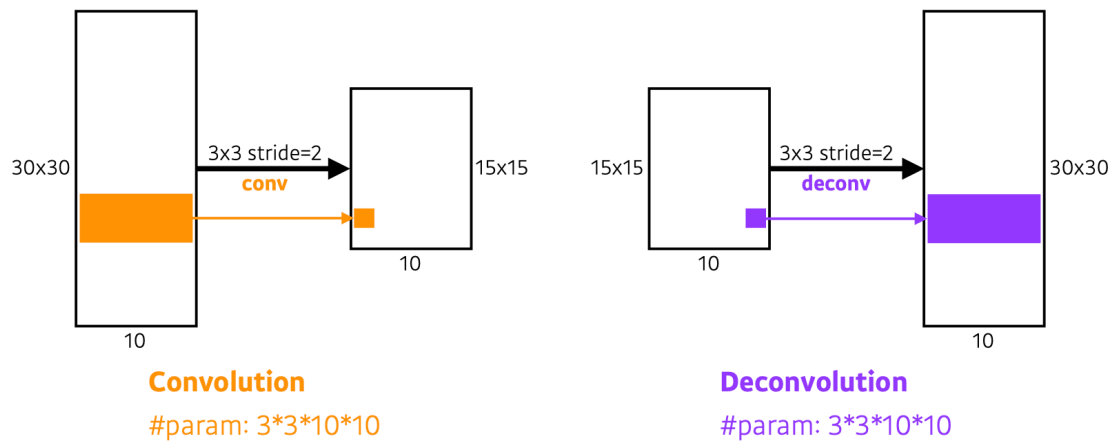
- Left: $4 \times 4 \times 16 \times 10 = 2,560$
- Right: $4 \times 4 \times 16 \times 10 = 2,560$

○ Deconvolution (Upsampling)



heatmap: 하나의 class 대표하는 coarse(대략적인) 정보

aim: dense prediction => **coarse map**을 **dense map**으로 복원, **deconvolution** 필요



- Convolution의 역연산 존재 X => deconvolution은 convolution의 역연산, 엄밀하게는 아님
But, 직관적인 이해를 위해서 역연산으로 생각함

○ Skip Architecture

coarse map을 dense map으로 복원 => 근본적으로 feature map 크기가 작음 => 여전히 coarse map

■ ZFNet 시각화

얕은 층에서는 주로 직선, 곡선, 생상, 등 local feature를

깊은 층에서는 얕은 층에 비해 개체의 일부분, 위치나 자세 변화 같은 global feature 추출

=> Deep & Coarse 레이어의 global feature & Shallow & Fine 층의 local feature 결합한 **skip connection** 이용