

Fully Convolutional Network

이지현 이채영

Contents

- What? Semantic Segmentation
- From Image Classification to Semantic Segmentation
- Upsampling via Deconvolution
- Skip Layer (Skip Connection / Skip combining)
- Fusing the Output
- Results

Semantic Segmentation

- Classify the object class for each Pixel within an image.
- That means there is a label for each pixel



segmented

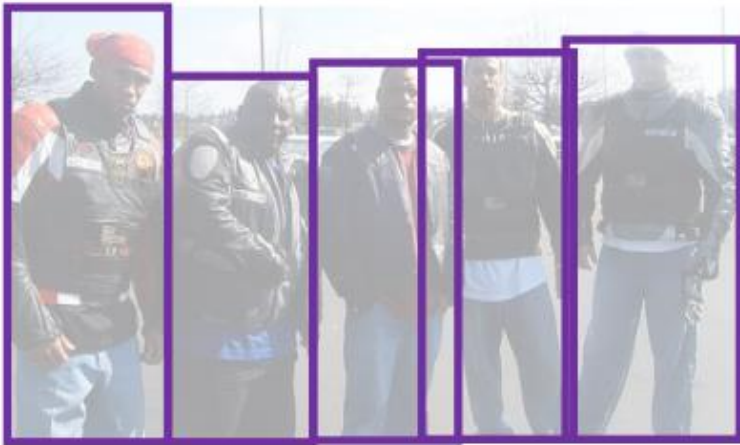
- 1: Person
- 2: Purse
- 3: Plants/Grass
- 4: Sidewalk
- 5: Building/Structures



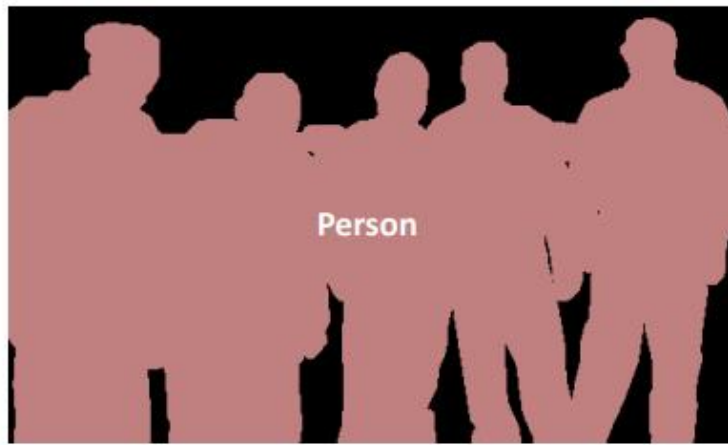
Segmentation map

Semantic Segmentation vs. Instance Segmentation

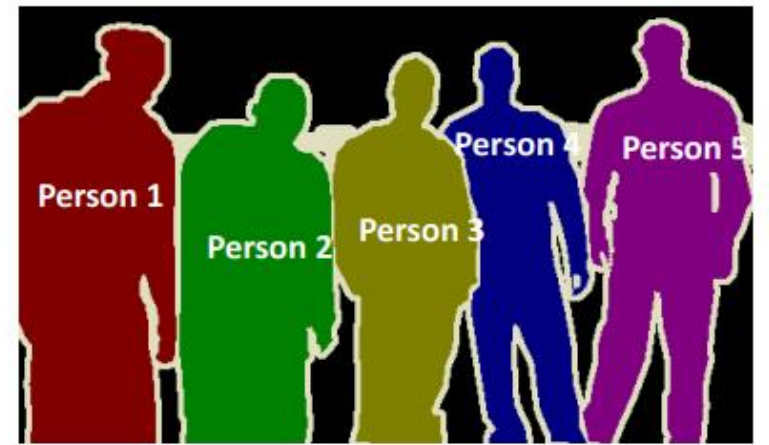
- **Object Detection:** detect the object in an image
- **Semantic Segmentation:** detect for each pixel, the object category it belongs to
- **Instance Segmentation:** identify for each pixel, the object instance it belongs to



Object Detection



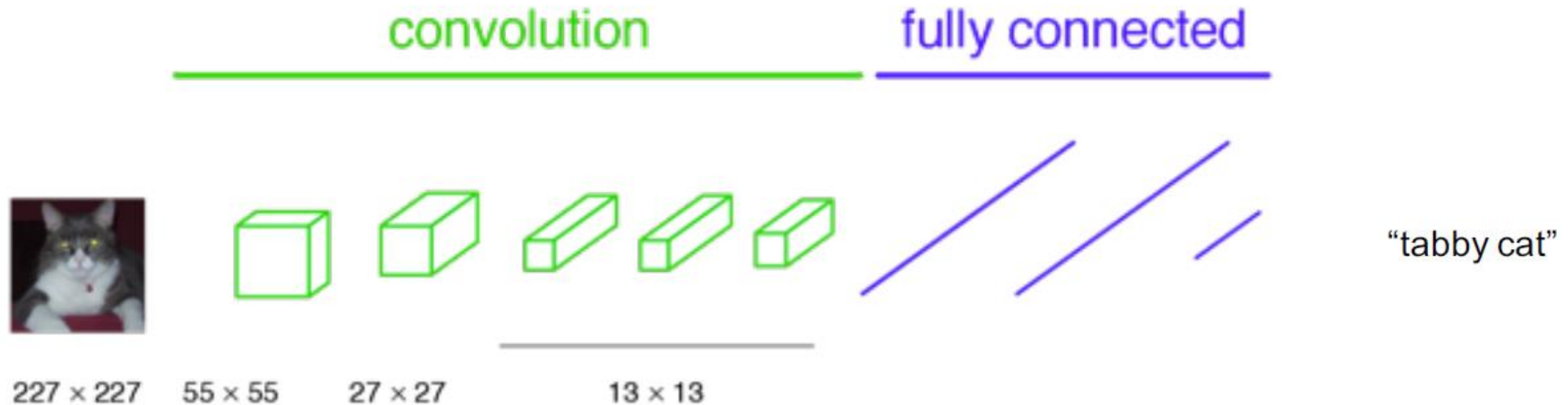
Semantic Segmentation



Instance Segmentation

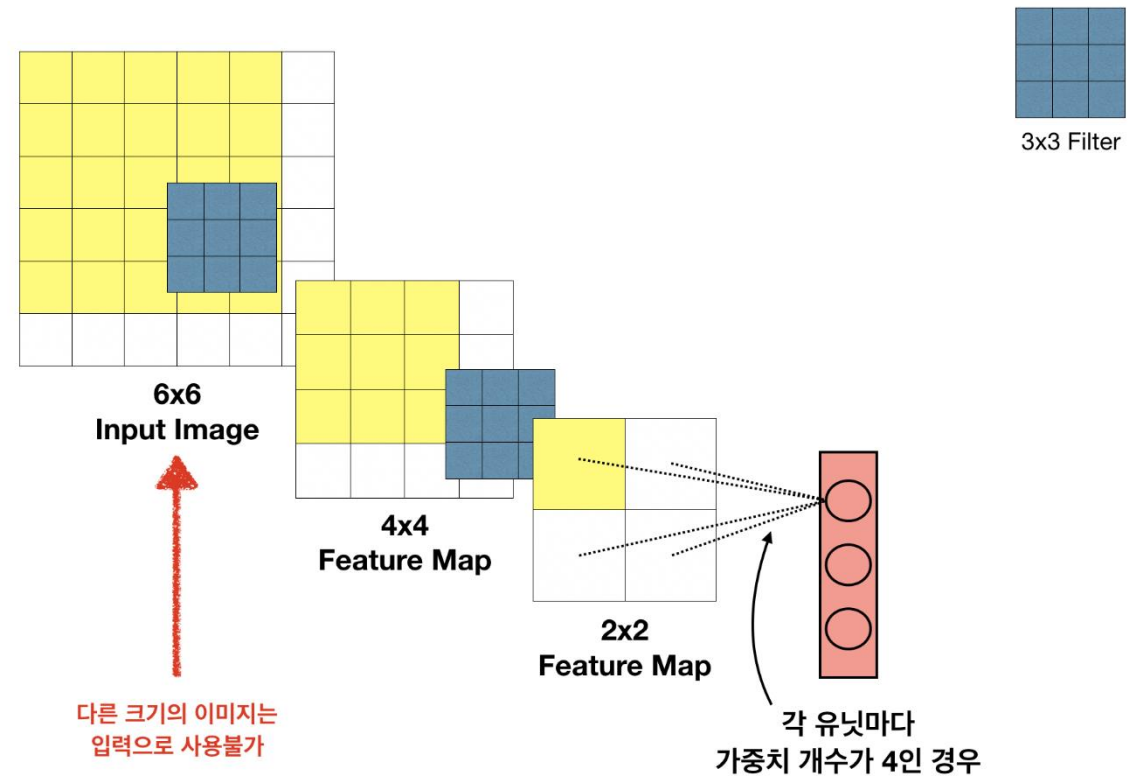
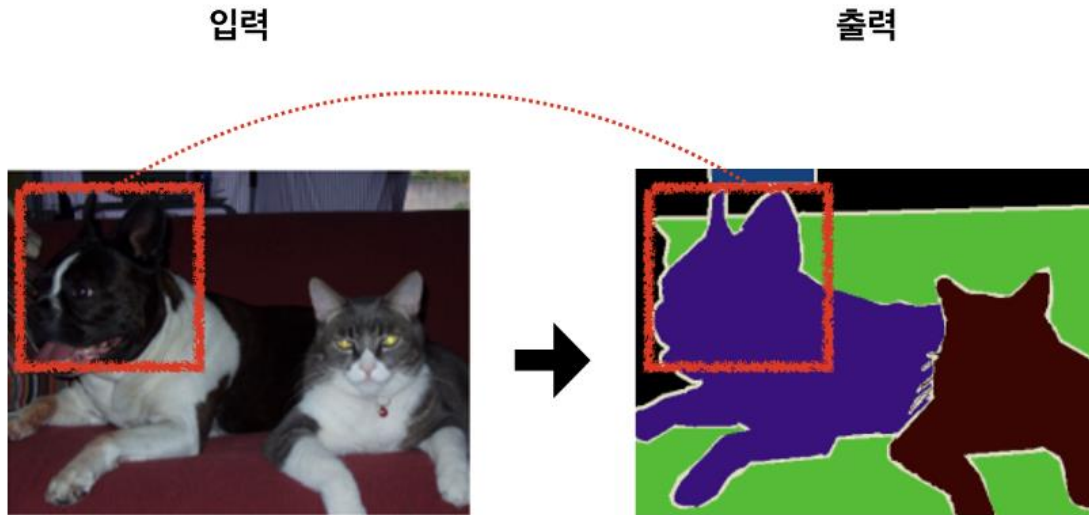
From Image Classification to Semantic Segmentation

- Image Classification



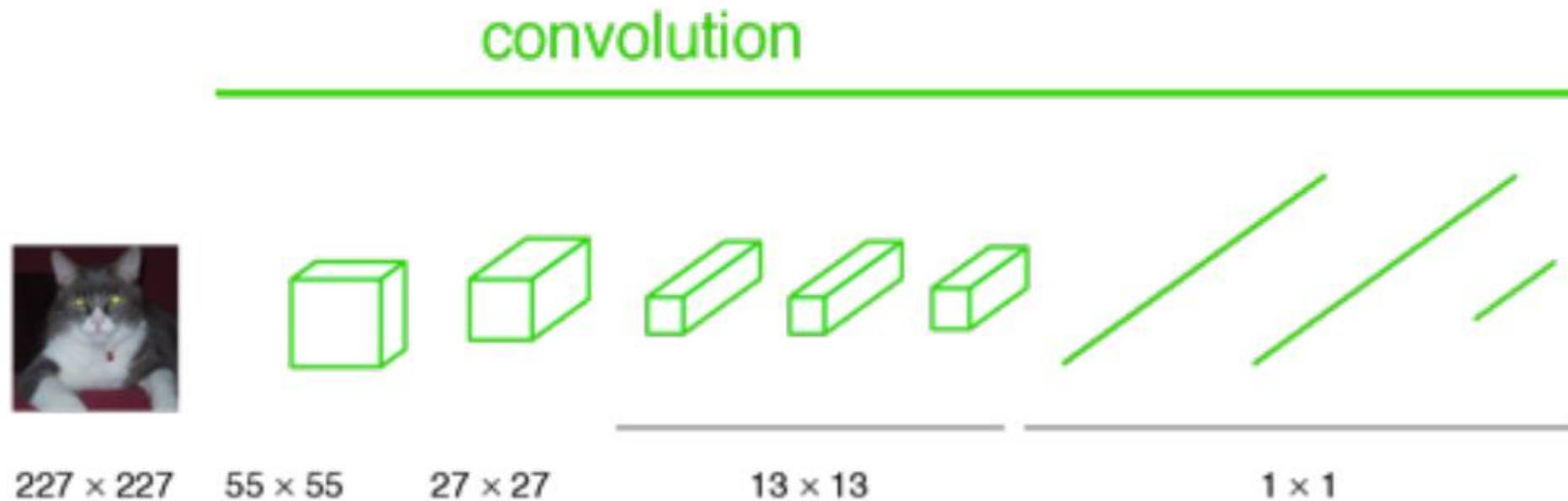
From Image Classification to Semantic Segmentation

- Image Segmentation



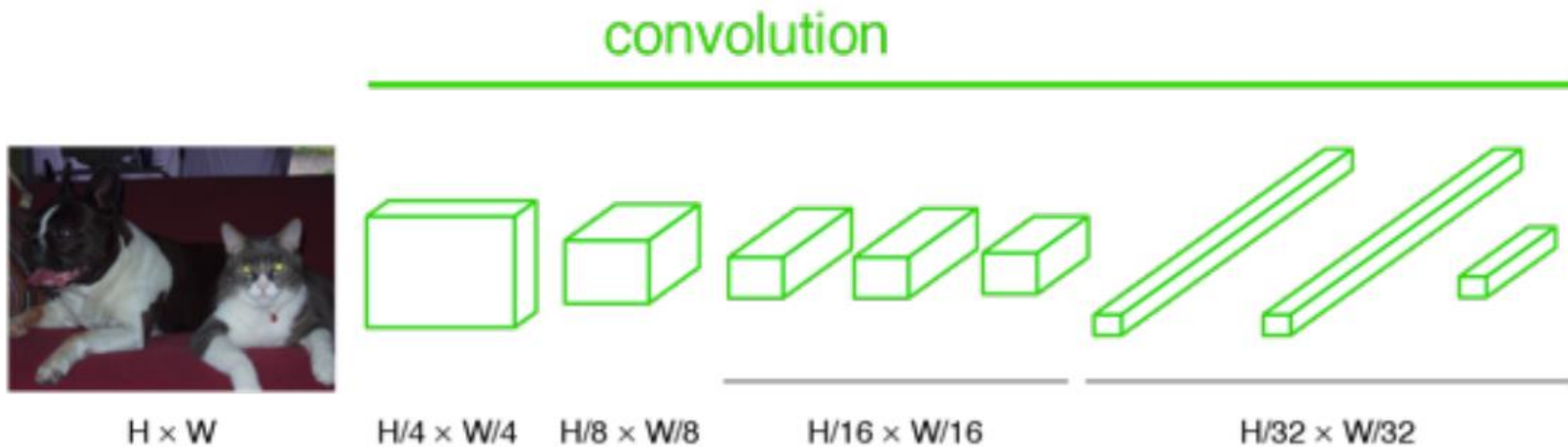
From Image Classification to Semantic Segmentation

- Convolutionalization



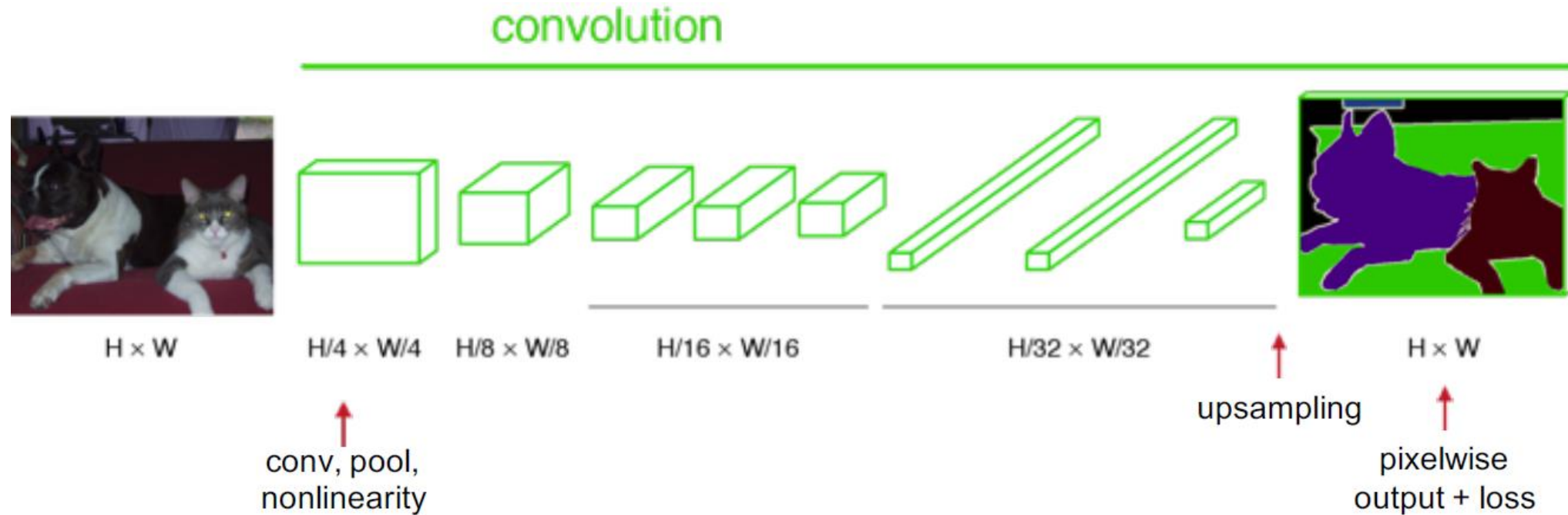
From Image Classification to Semantic Segmentation

- Convolutionalization



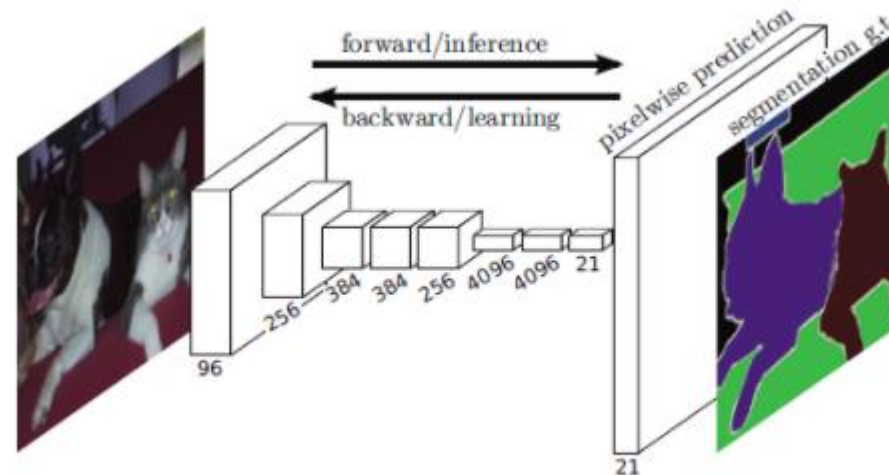
From Image Classification to Semantic Segmentation

- Convolutionalization



Upsampling via Deconvolution

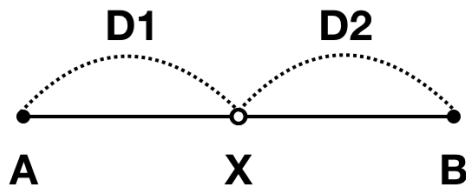
- Convolutionalization의 문제점 보완
 - Linear Interpolation (선형보간법)
 - Unpooling
 - Transposed convolution (= Deconvolution)



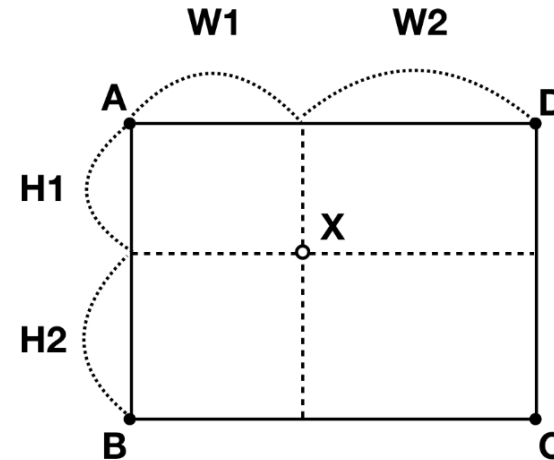
Feature Map / Filter Number Along Layers

Upsampling Via Deconvolution

- **Linear Interpolation & Bilinear Interpolation**



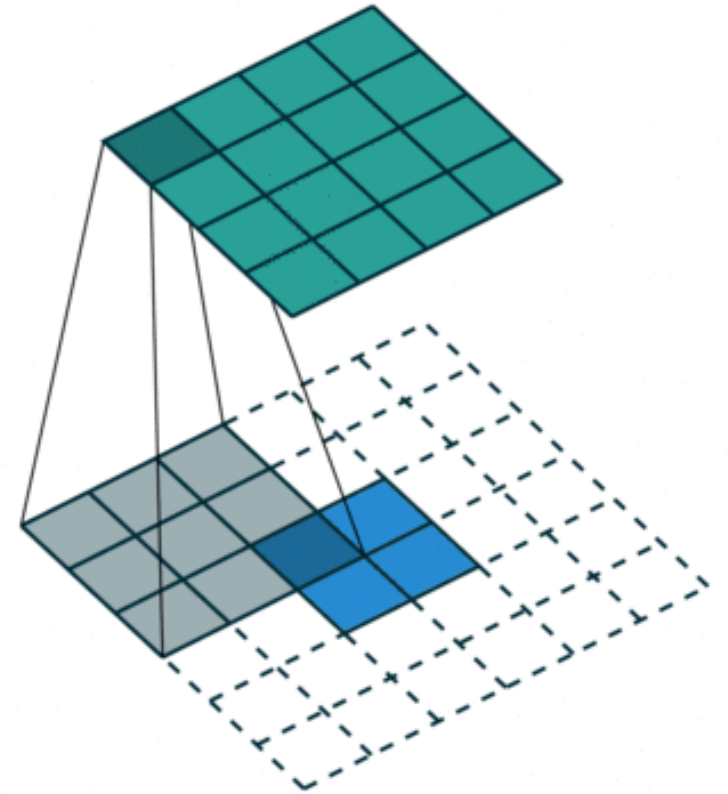
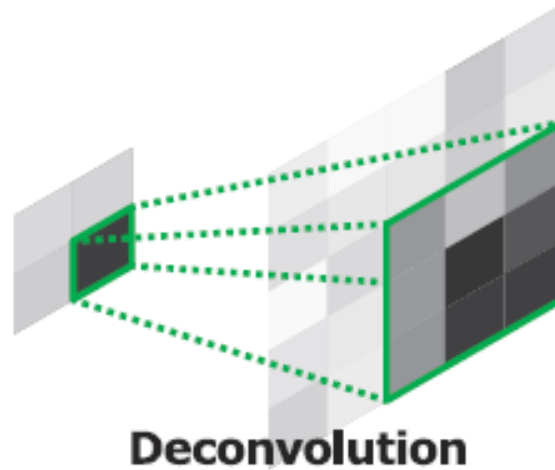
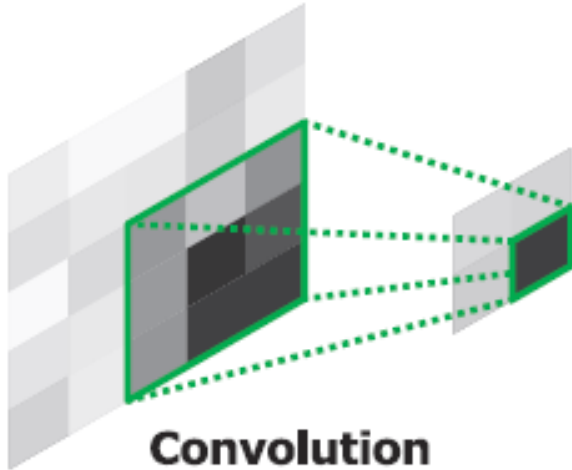
$$A \frac{D2}{D1 + D2} + B \frac{D1}{D1 + D2}$$



$$X = \left(A \frac{H2}{H1 + H2} + B \frac{H1}{H1 + H2} \right) \frac{W2}{W1 + W2} + \left(D \frac{H2}{H1 + H2} + C \frac{H1}{H1 + H2} \right) \frac{W1}{W1 + W2}$$

Upsampling Via Deconvolution

- Deconvolution



Upsampling Via Deconvolution

- Unpooling

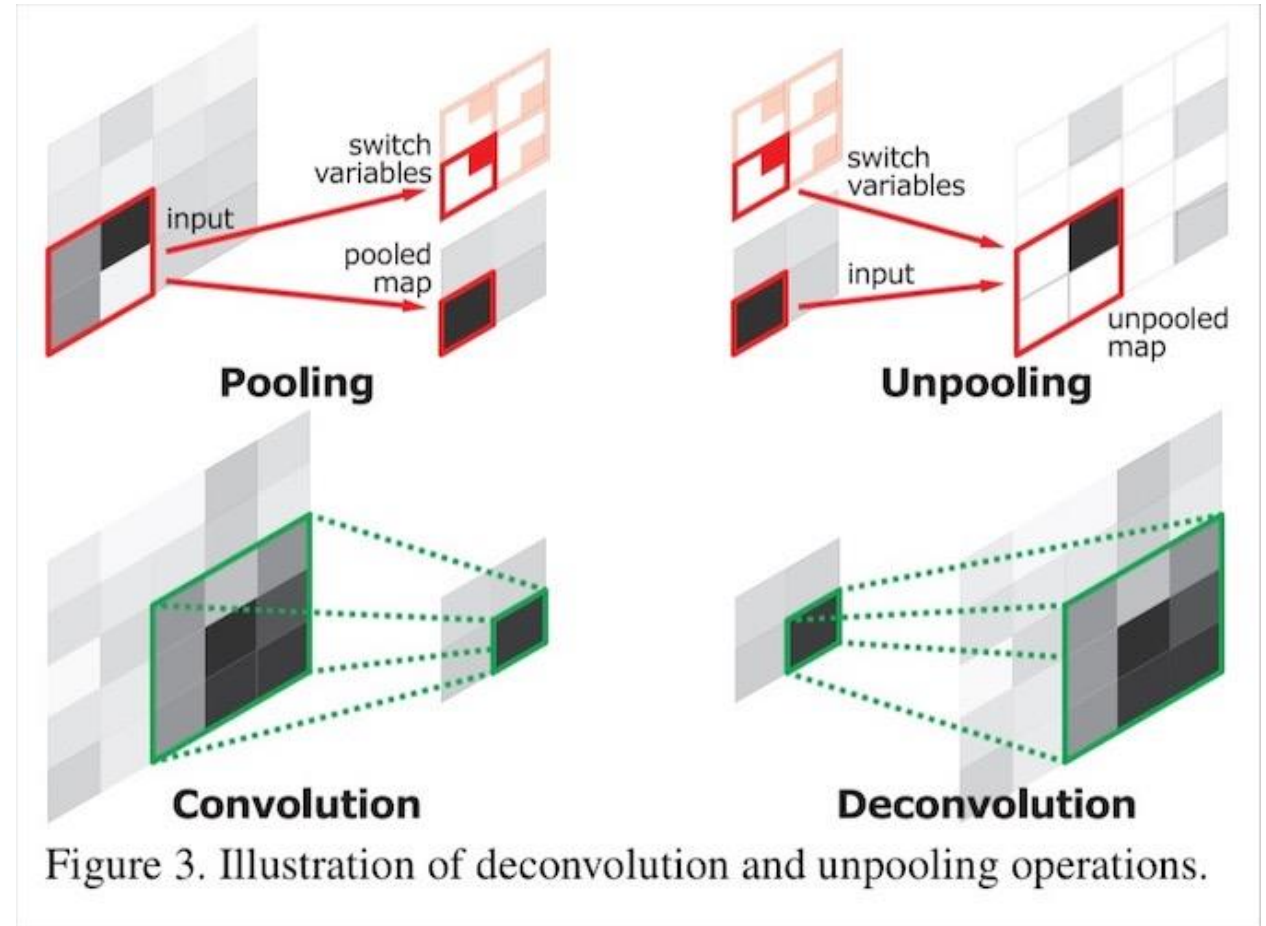
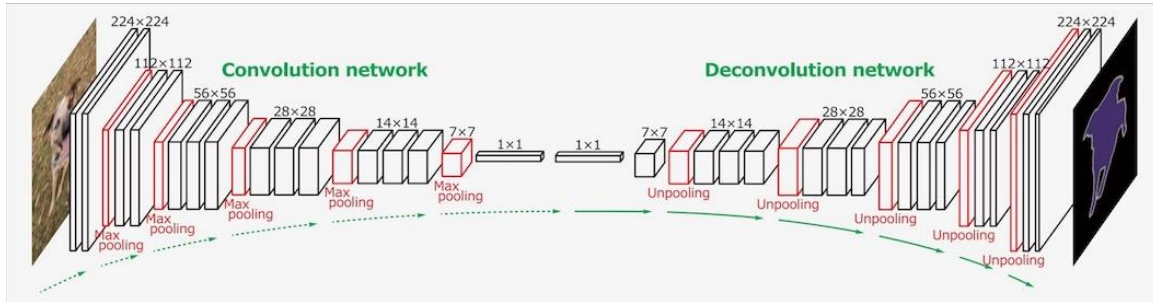
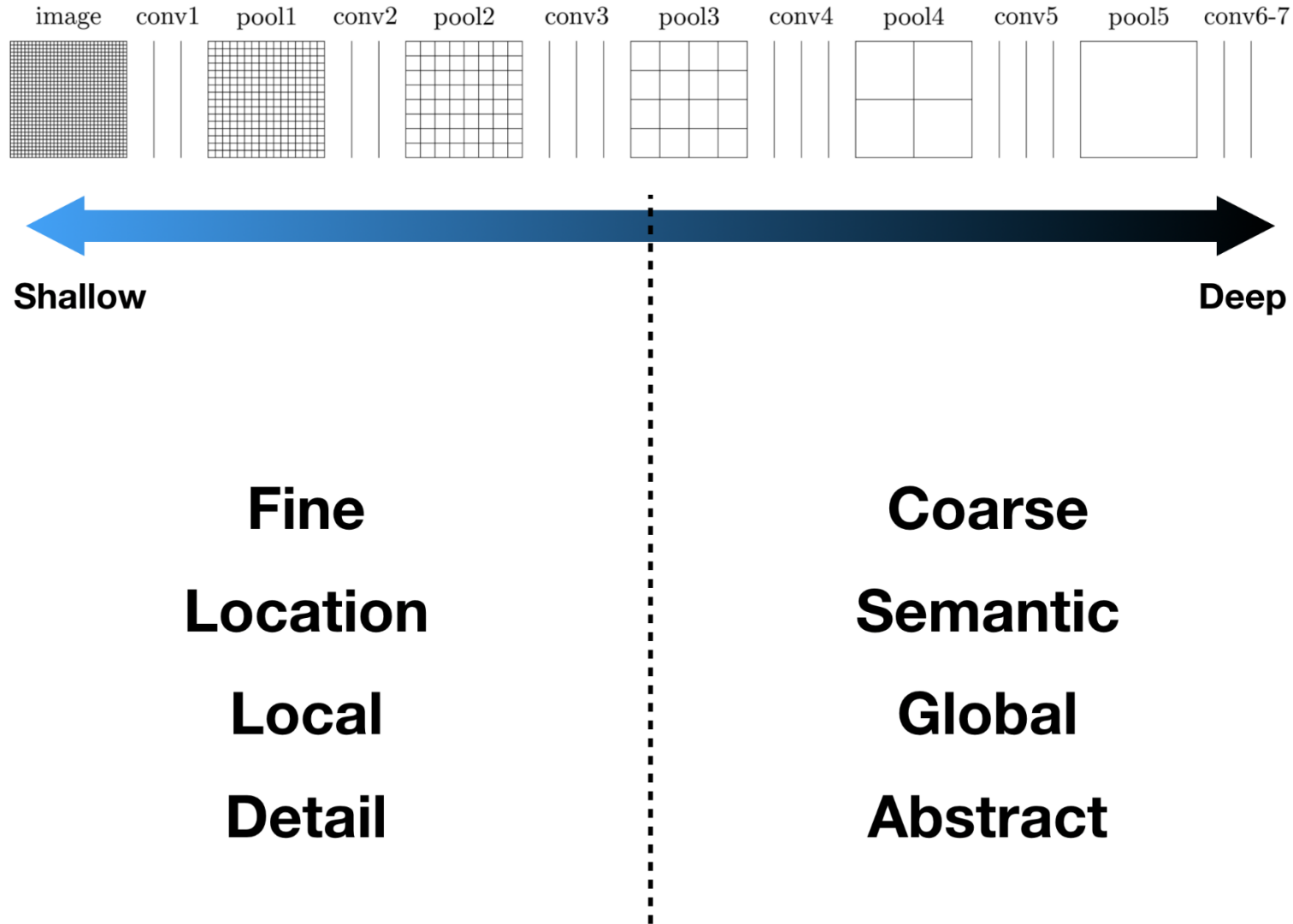


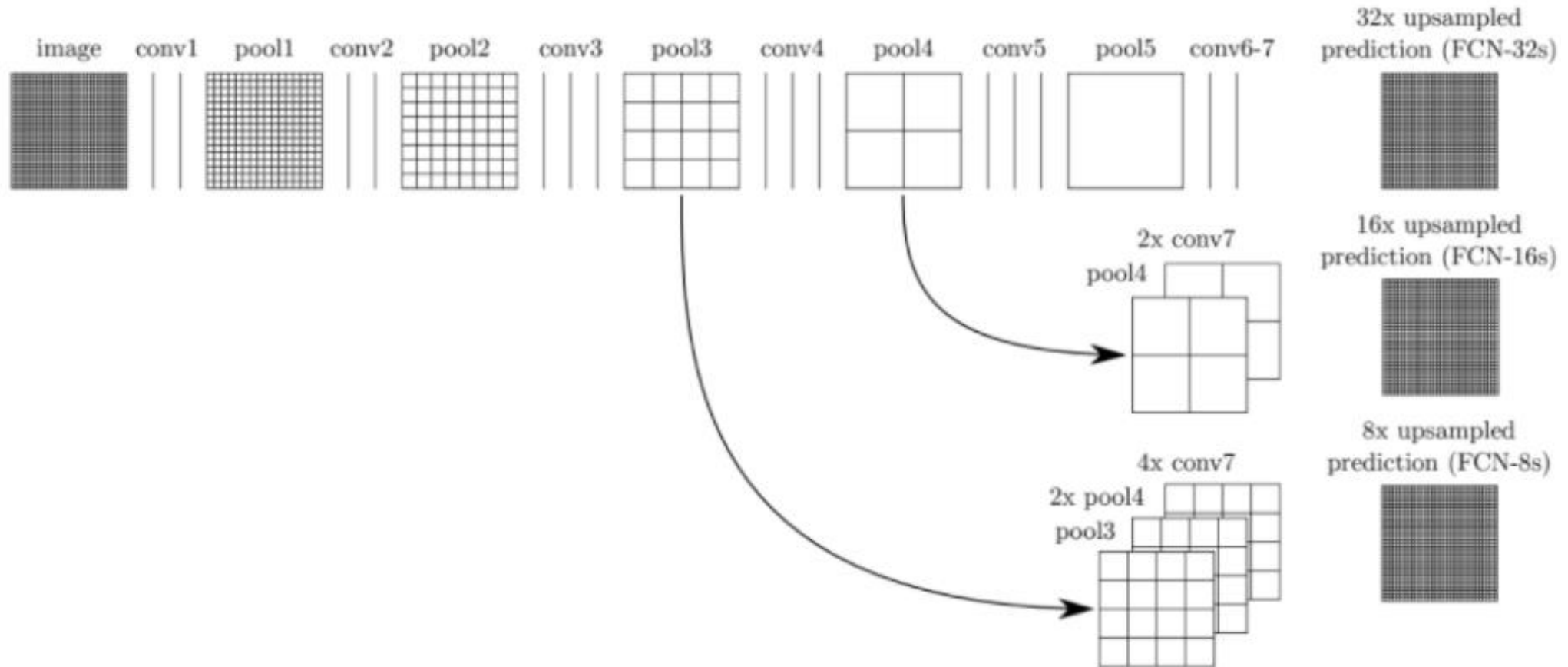
Figure 3. Illustration of deconvolution and unpooling operations.

Skip Layer (Skip Connection / Skip combining)



Skip Layer (Skip Connection / Skip combining)

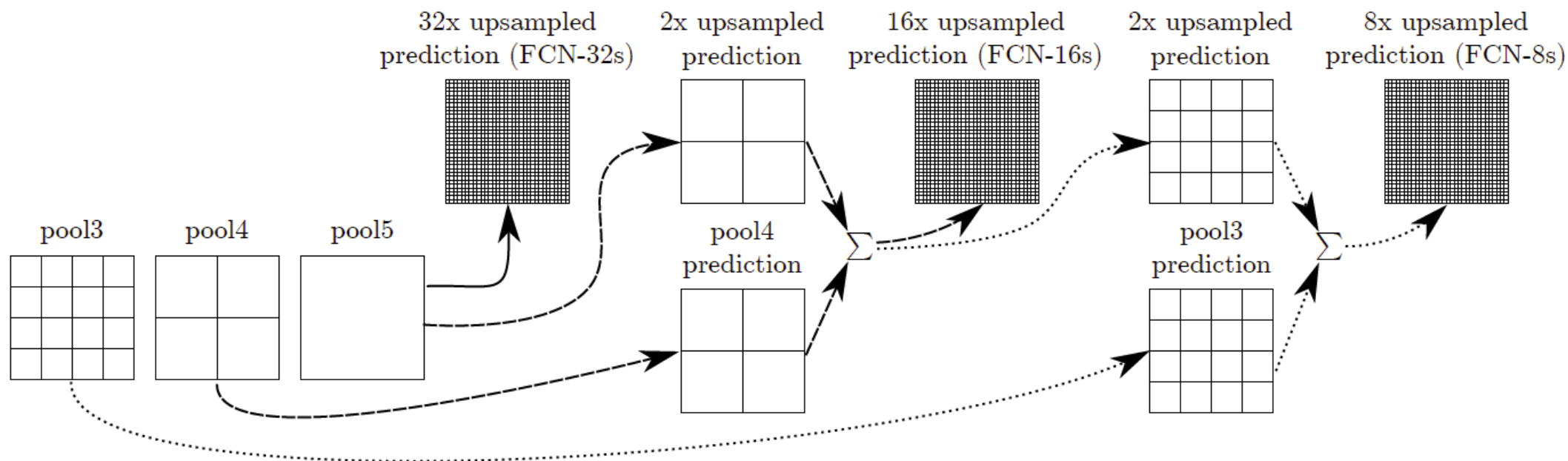
- Skip layer의 원리



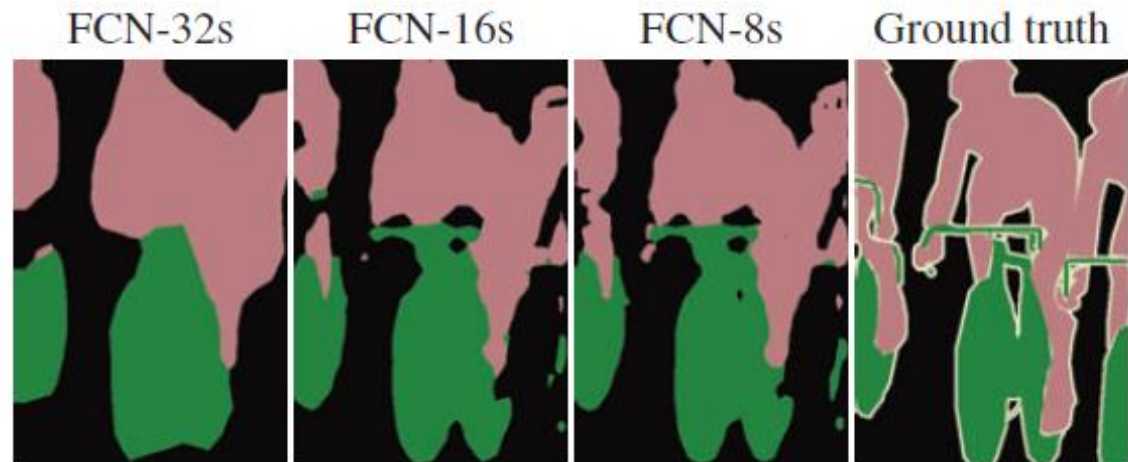
Fusing the Output

- Segmentation

: 각 class 별로 upsampling 된 결과를 사용



Results



	pixel acc.	mean acc.	mean IU	f.w. IU		pixel acc.	mean acc.	mean IU	f.w. IU		pixel acc.	mean acc.	mean IU	f.w. IU	geom. acc.
FCN-32s-fixed	83.0	59.7	45.4	72.0	Gupta <i>et al.</i> [14]	60.3	-	28.6	47.0	Liu <i>et al.</i> [23]	76.7	-	-	-	-
FCN-32s	89.1	73.3	59.4	81.4	FCN-32s RGB	60.0	42.2	29.2	43.9	Tighe <i>et al.</i> [33]	-	-	-	-	90.8
FCN-16s	90.0	75.7	62.4	83.0	FCN-32s RGBD	61.5	42.4	30.5	45.5	Tighe <i>et al.</i> [34] 1	75.6	41.1	-	-	-
FCN-8s	90.3	75.9	62.7	83.2	FCN-32s HHA	57.1	35.2	24.2	40.4	Tighe <i>et al.</i> [34] 2	78.6	39.2	-	-	-
					FCN-32s RGB-HHA	64.3	44.9	32.8	48.0	Farabet <i>et al.</i> [8] 1	72.3	50.8	-	-	-
					FCN-16s RGB-HHA	65.4	46.1	34.0	49.5	Farabet <i>et al.</i> [8] 2	78.5	29.6	-	-	-
										Pinheiro <i>et al.</i> [28]	77.7	29.8	-	-	-
										FCN-16s	85.2	51.7	39.5	76.1	94.3

Thank you!
