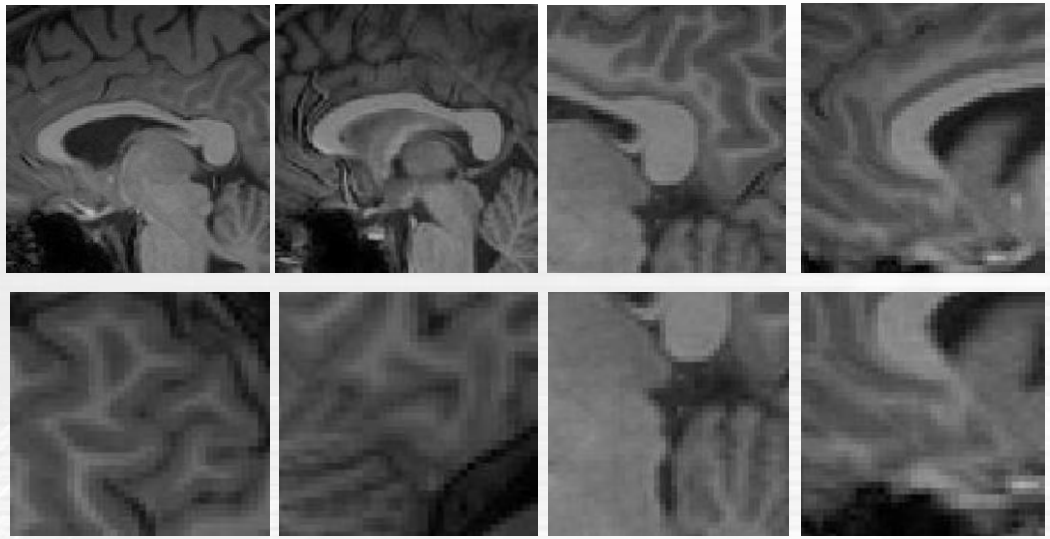
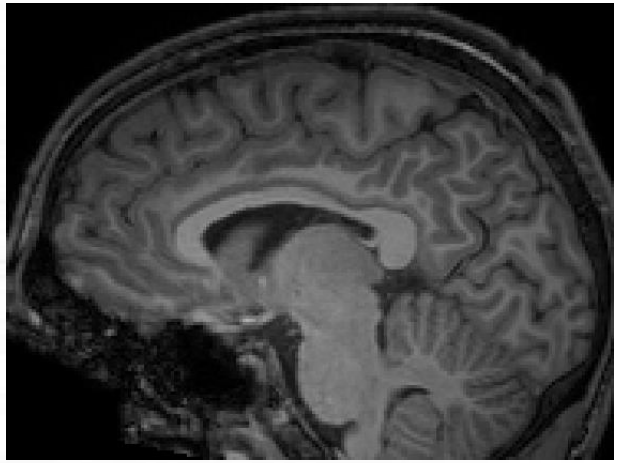
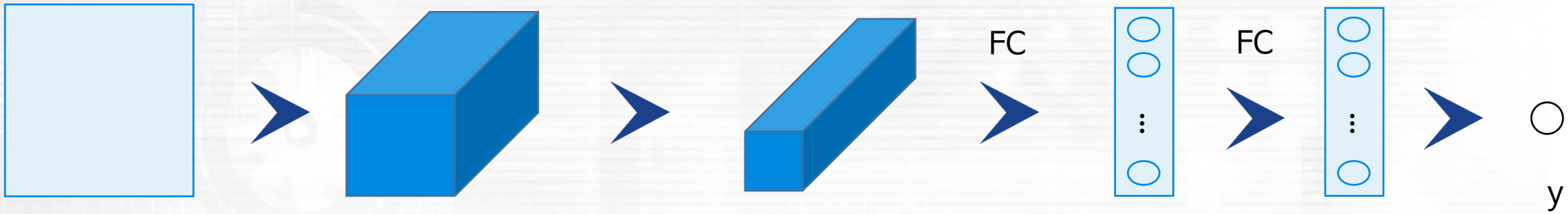
Segmentation using Deep Learning

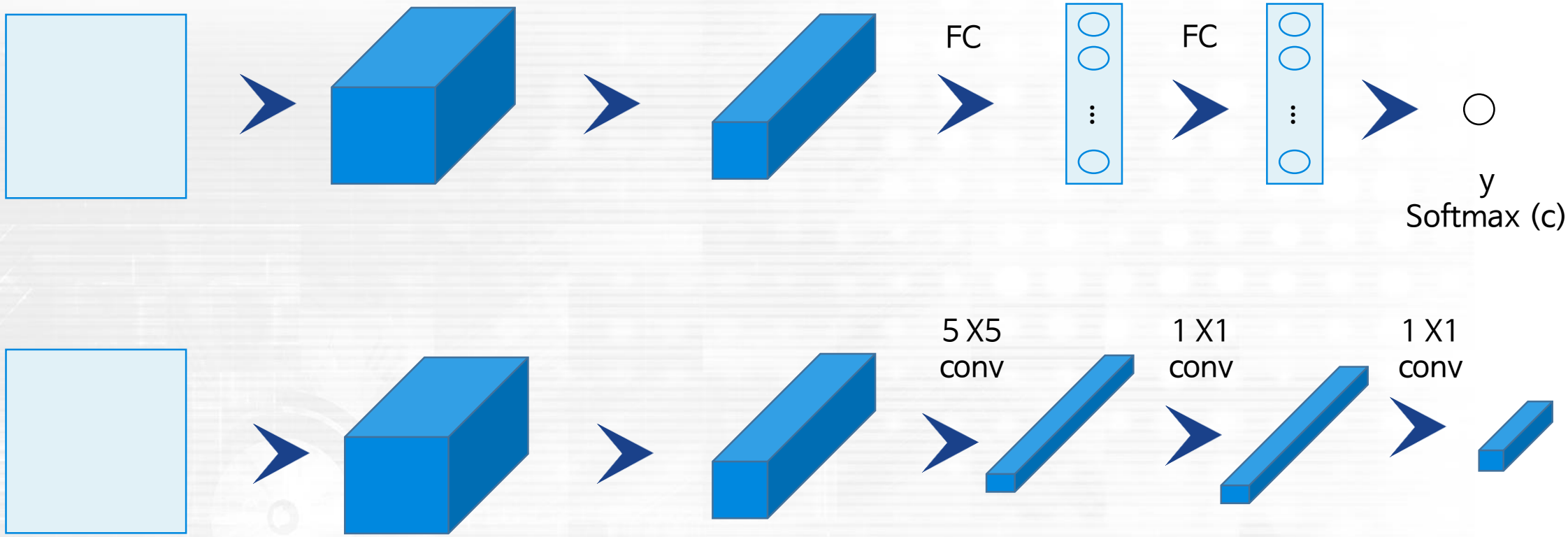
Segmentation using Deep Learning



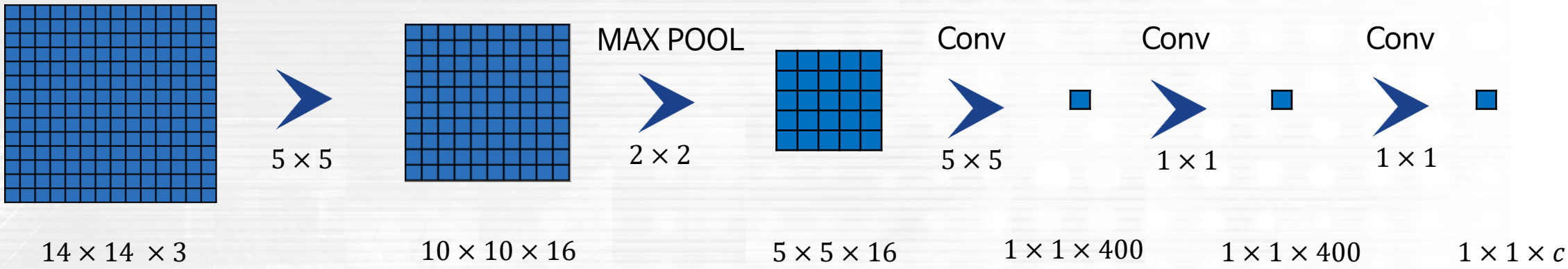




Fully Convolutional Networks

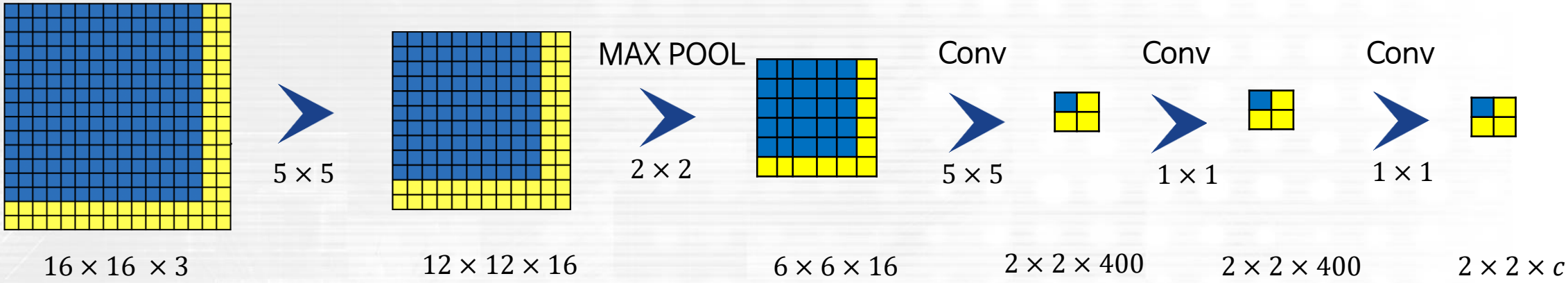


Fully Convolutional Networks



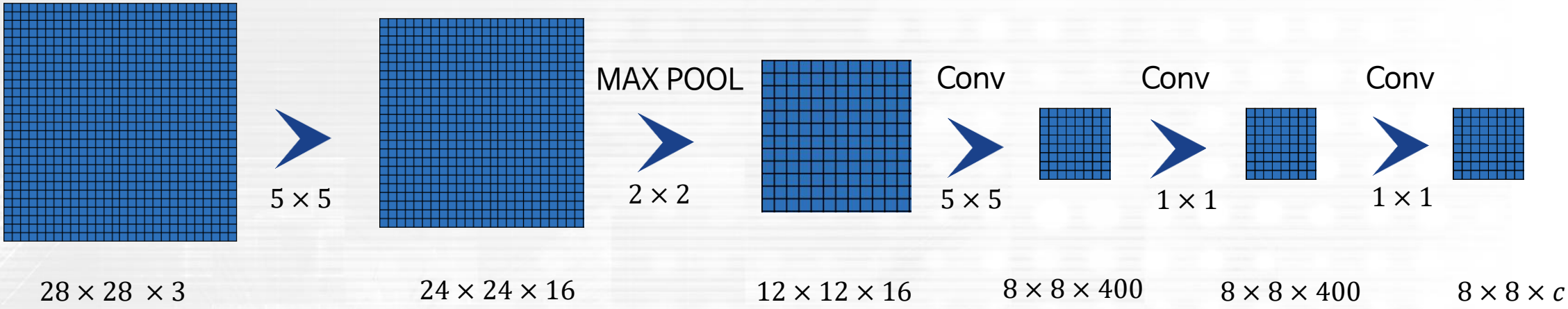
출처 [Sermanet et al., 2014, OverFeat: Integrated recognition, localization and detection using convolutional networks]

Fully Convolutional Networks



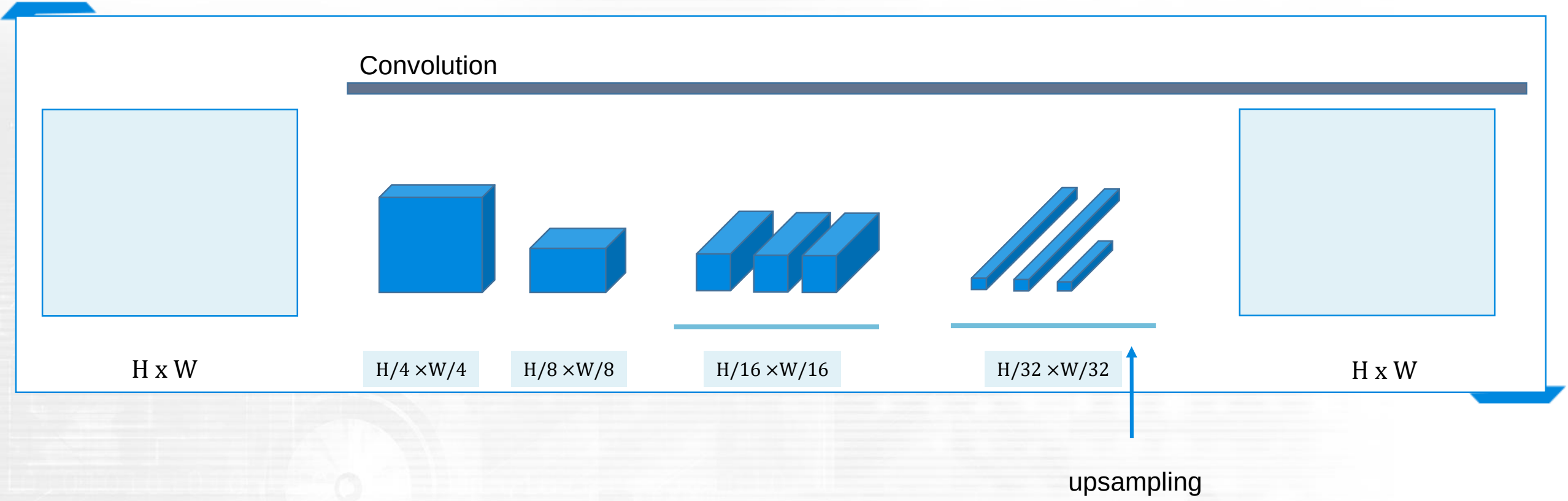
출처 [Sermanet et al., 2014, OverFeat: Integrated recognition, localization and detection using convolutional networks]

Fully Convolutional Networks



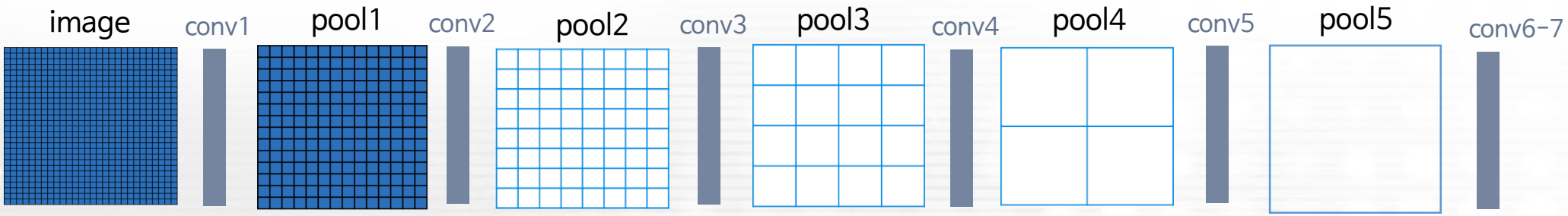
출처 [Sermanet et al., 2014, OverFeat: Integrated recognition, localization and detection using convolutional networks]

Fully Convolutional Networks



출처 Jonathan Long et al. Fully Convolutional Networks for Semantic Segmentation, CVPR 2015

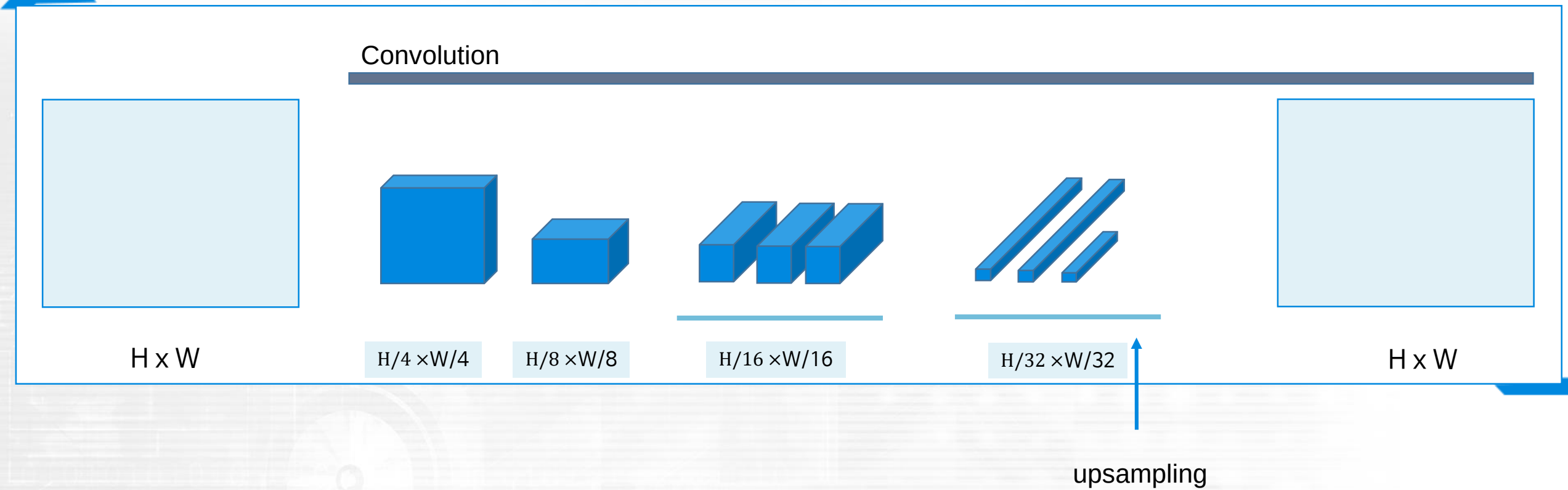
Fully Convolutional Networks



출처 Jonathan Long et al. Fully Convolutional Networks for Semantic Segmentation, CVPR 2015

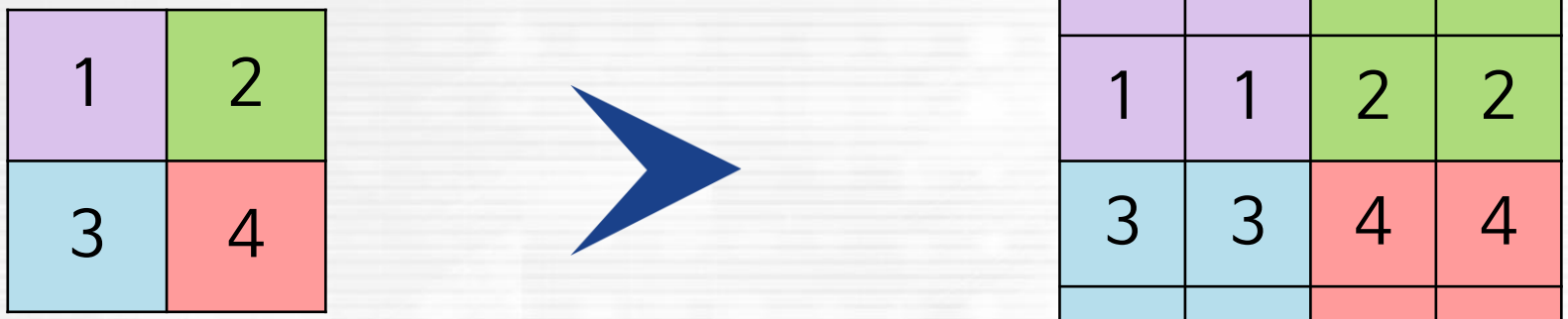
Fully Convolutional Networks

Fully Convolutional Networks



Upsampling

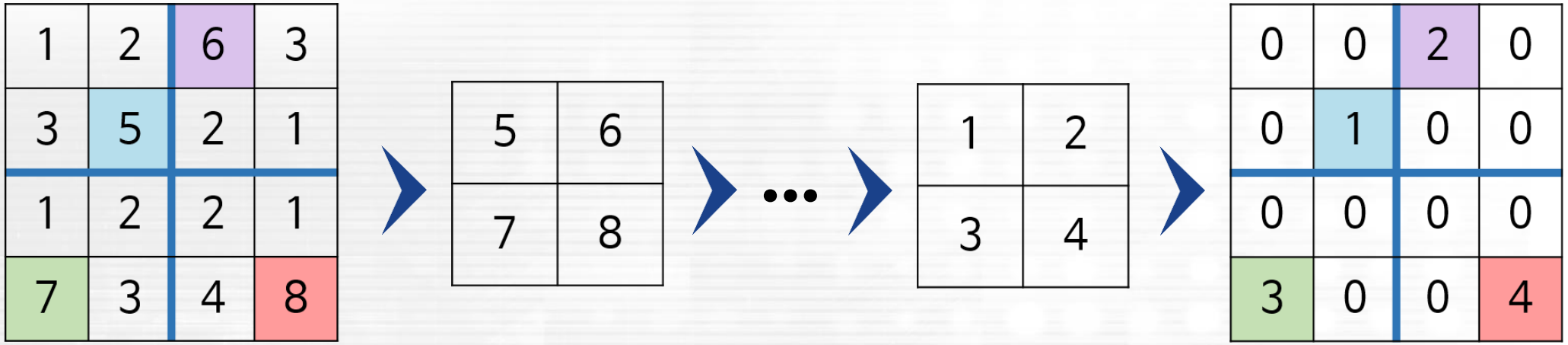
Nearest Neighbor

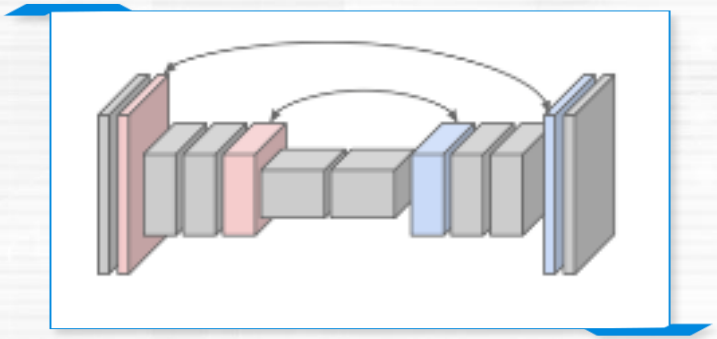


Input : 2 x 2

Output : 4 x 4

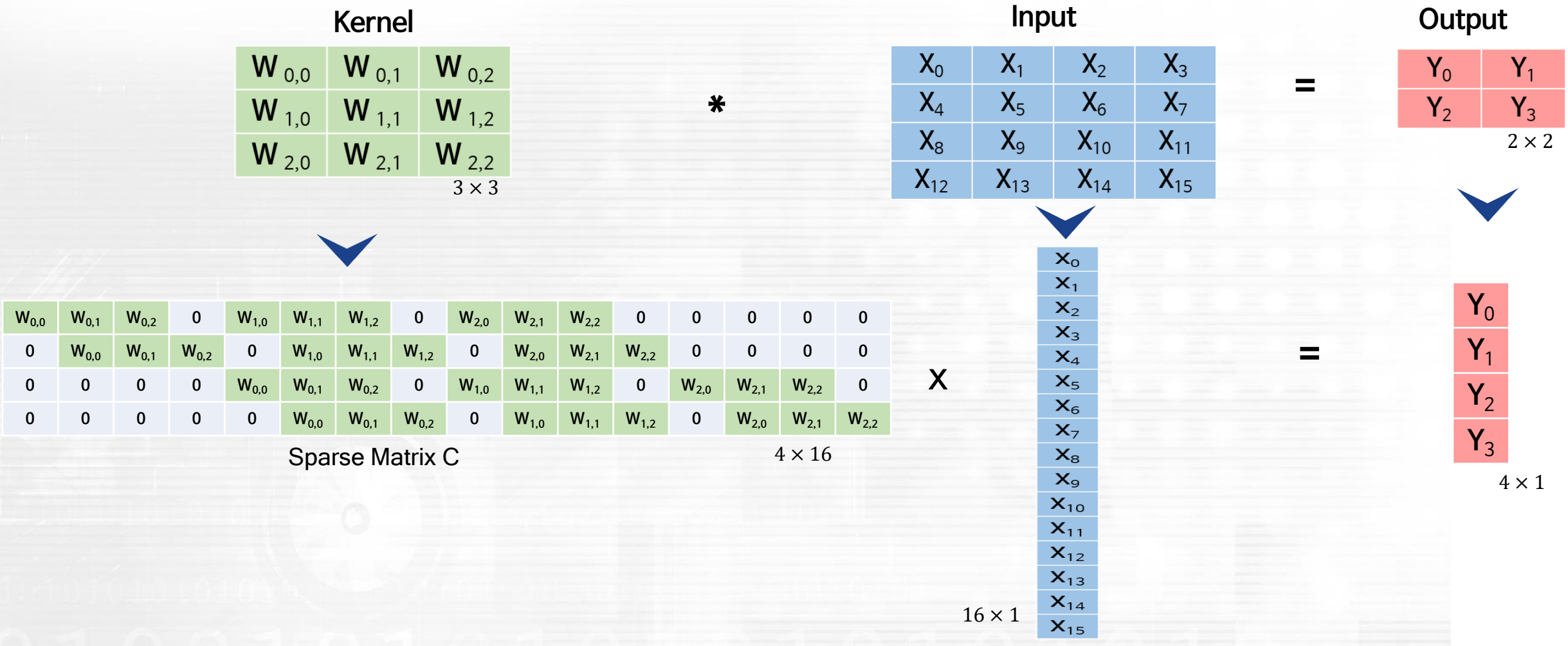
Unpooling





출처 Learning Deconvolution Network for Semantic Segmentation, ICCV 2015

Transposed Convolution



Transposed Convolution

$W_{0,0}$	0	0	0
$W_{0,1}$	$W_{0,0}$	0	0
$W_{0,2}$	$W_{0,1}$	0	0
0	$W_{0,2}$	0	0
$W_{1,0}$	0	$W_{0,0}$	0
$W_{1,1}$	$W_{1,0}$	$W_{0,1}$	$W_{0,0}$
$W_{1,2}$	$W_{1,1}$	$W_{0,2}$	$W_{0,1}$
	$W_{1,2}$	0	$W_{0,2}$
$W_{2,0}$	0	$W_{1,0}$	0
$W_{2,1}$	$W_{2,0}$	$W_{1,1}$	$W_{1,0}$
$W_{2,2}$	$W_{2,1}$	$W_{1,2}$	$W_{1,1}$
0	$W_{2,2}$	0	$W_{1,2}$
0	0	$W_{2,0}$	0
0	0	$W_{2,1}$	$W_{2,0}$
0	0	$W_{2,2}$	$W_{2,1}$
0	0	0	$W_{2,2}$

16×4

Sparse Matrix C^T

X

Y_0
Y_1
Y_2
Y_3

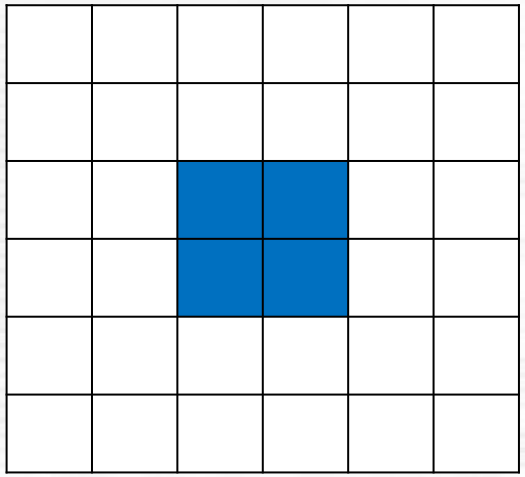
4×1

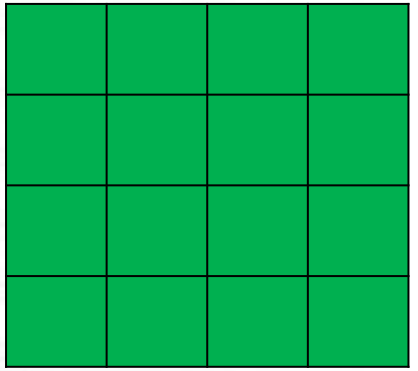
=

X_0
X_1
X_2
X_3
X_4
X_5
X_6
X_7
X_8
X_9
X_{10}
X_{11}
X_{12}
X_{13}
X_{14}
X_{15}

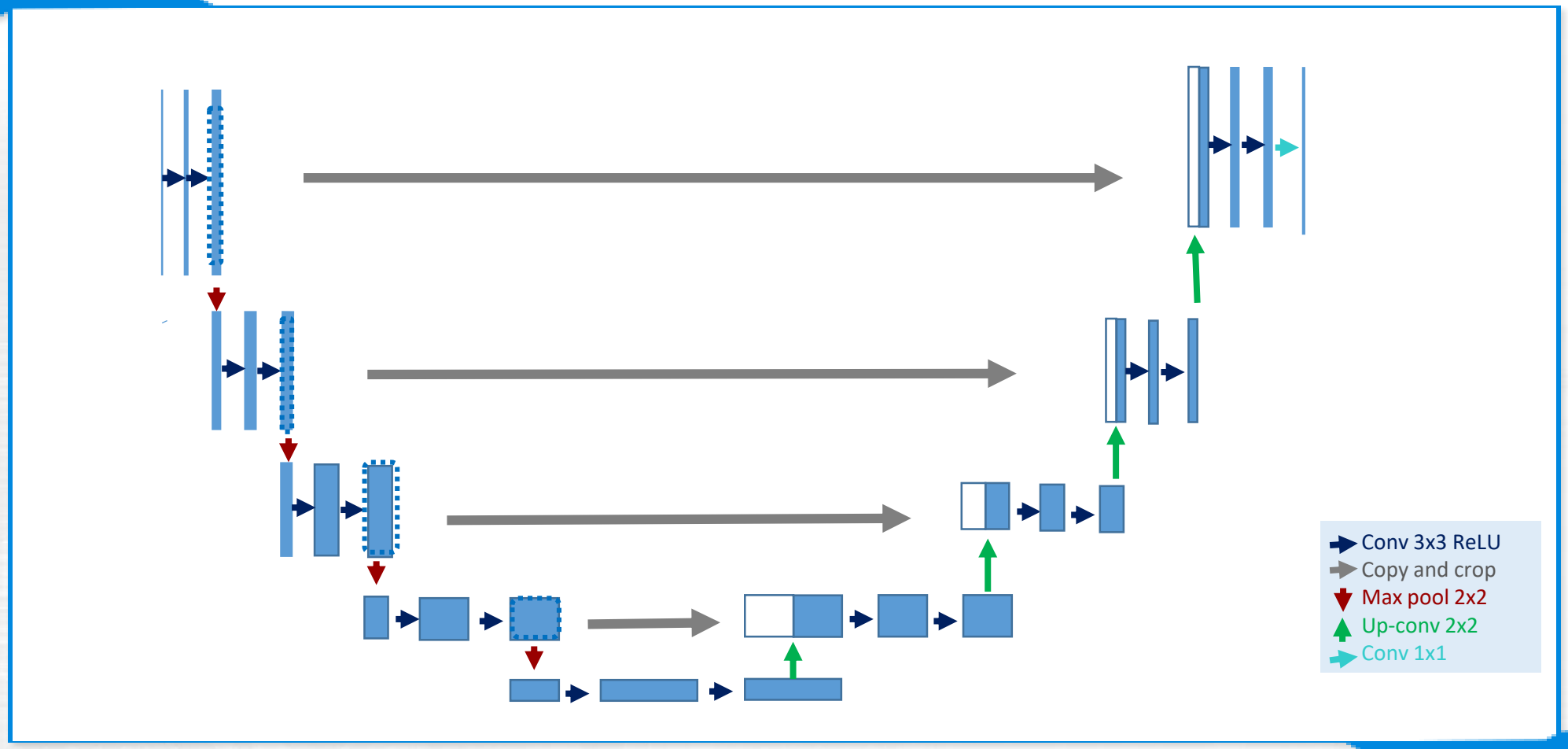
16×1

Up – Convolution



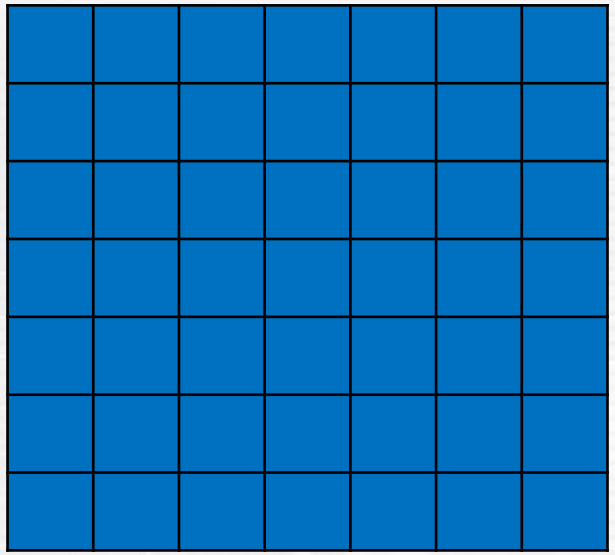


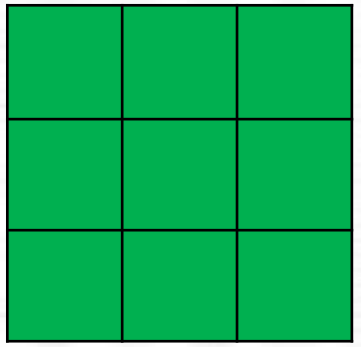
U-Net



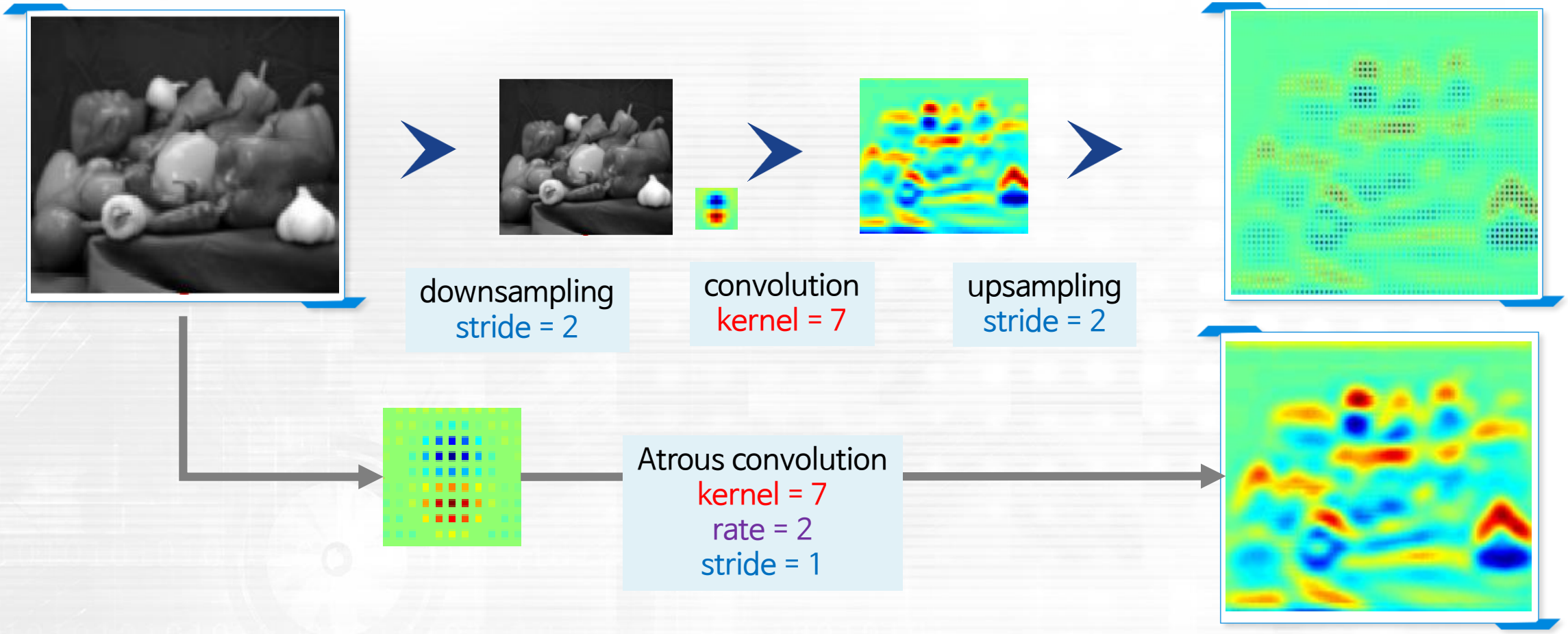
출처 U-Net: Convolutional Networks for Biomedical Image Segmentation, MICCAI 2015

Dilated Convolution



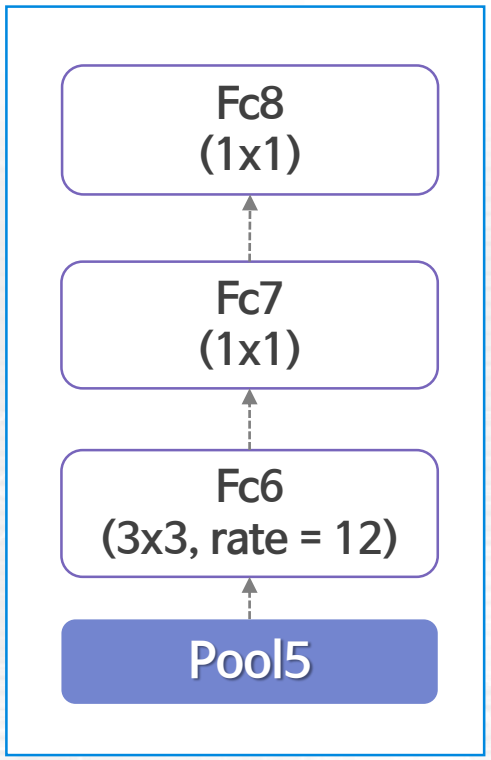


Dilated Convolution

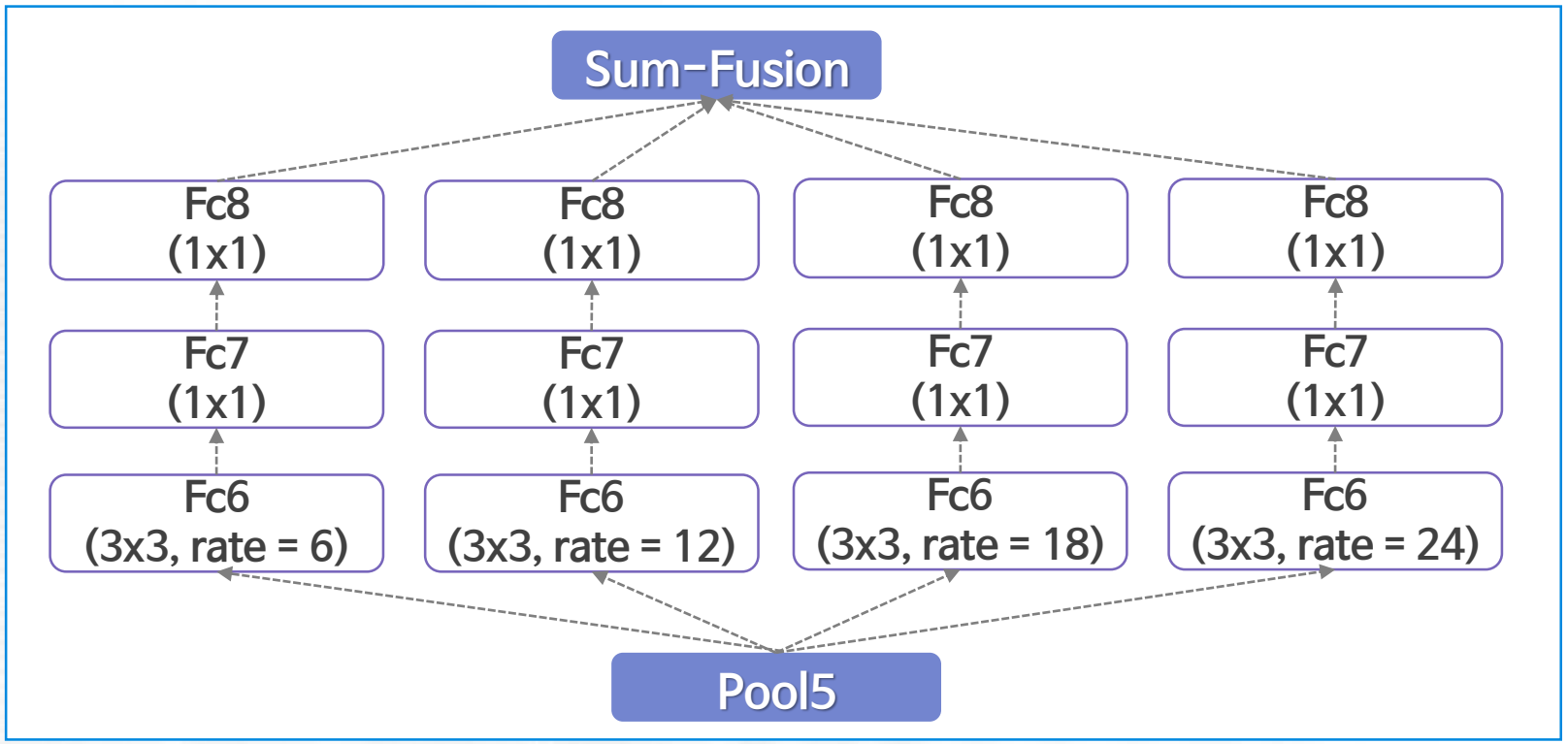


출처 DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs, TPAMI 2018

Atrous Spatial Pyramid Pooling



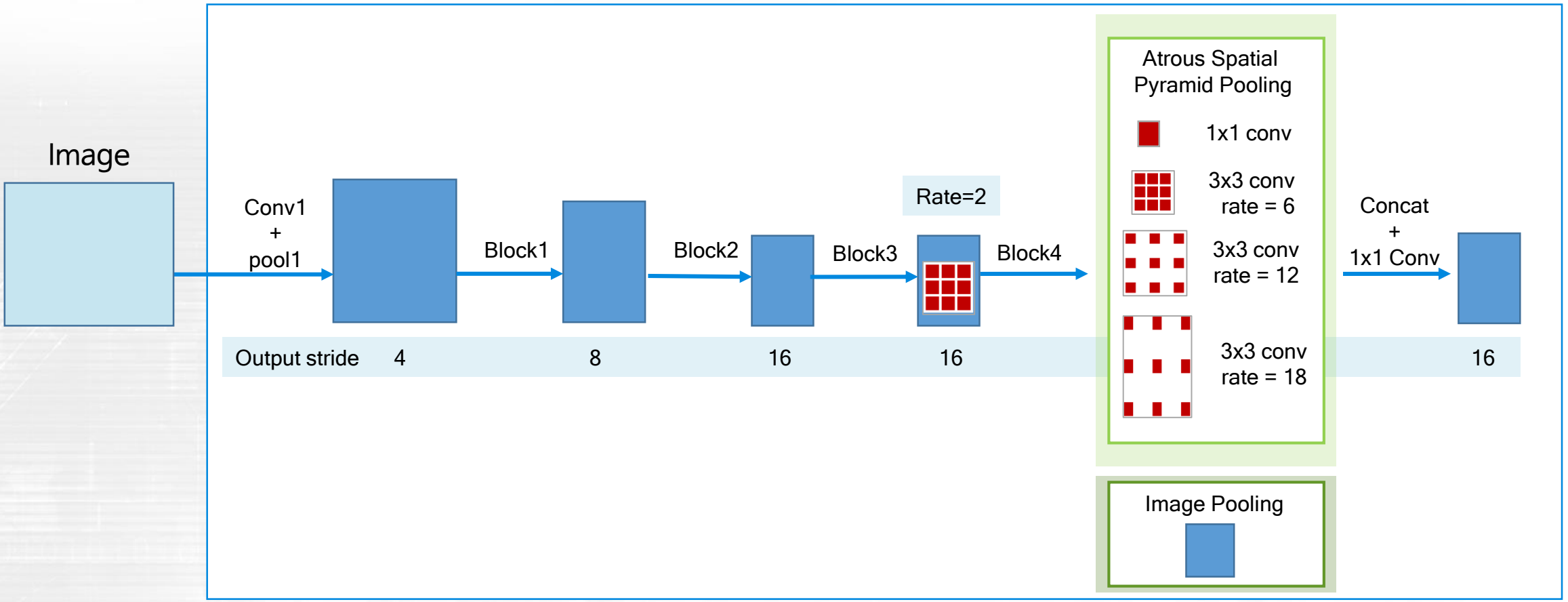
DeepLab-LargeFOV



DeepLab-ASPP

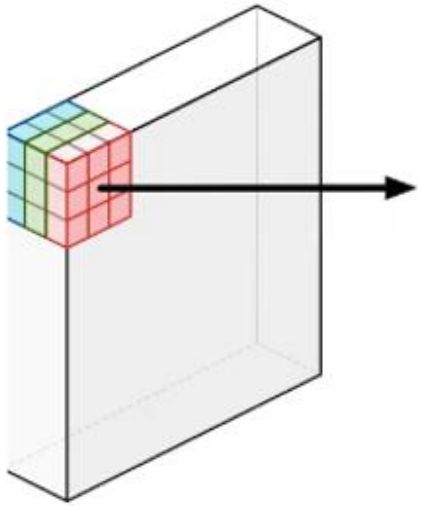
출처 DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs, TPAMI 2018

DeepLab V3



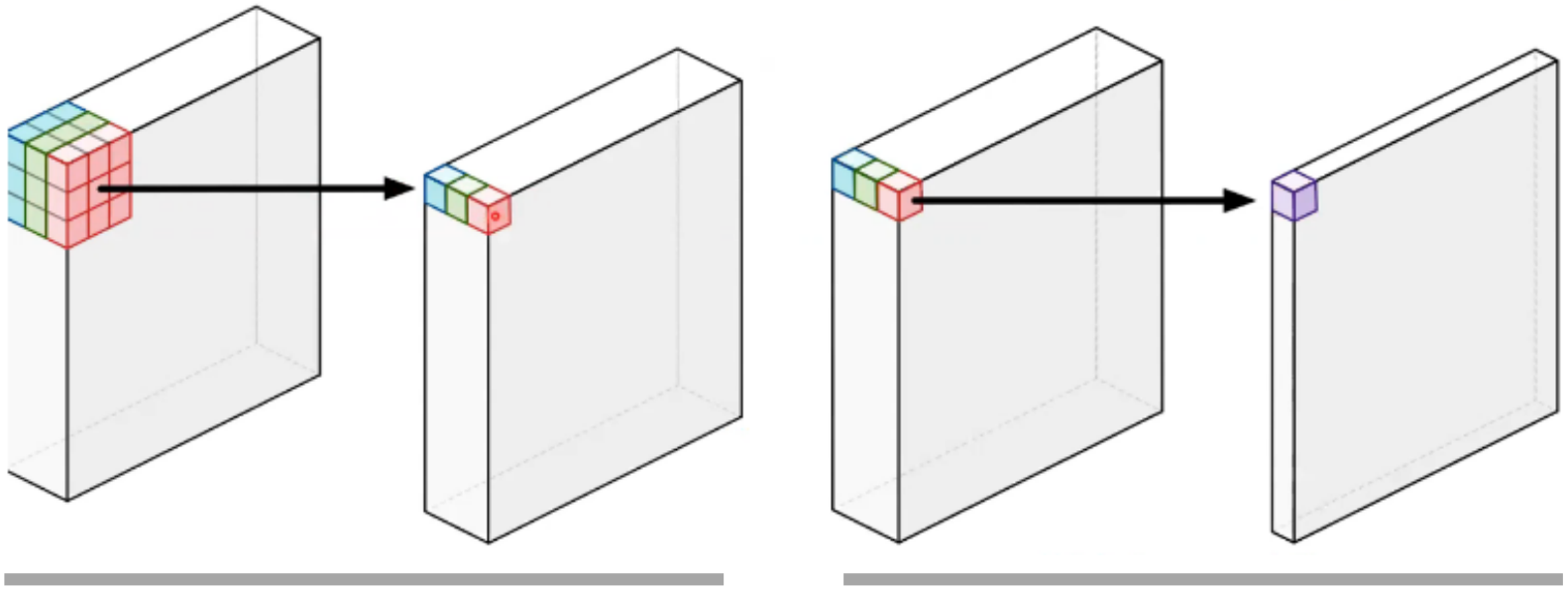
출처 Rethinking Atrous Convolution for Semantic Image Segmentation. arXiv 2017.
Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation. arXiv 2018.

Depth-wise Separable Convolution



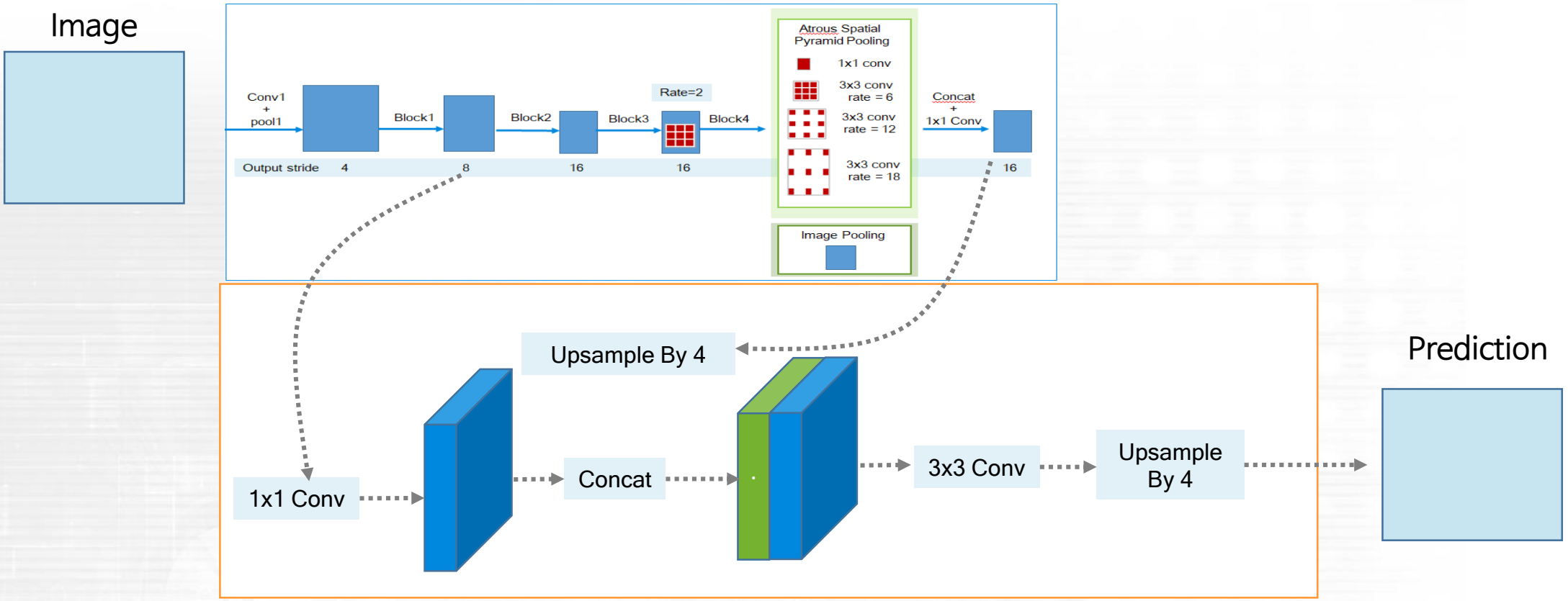
출처 Rethinking Atrous Convolution for Semantic Image Segmentation. arXiv 2017.
Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation. arXiv 2018.

Depth-wise Separable Convolution



출처 Rethinking Atrous Convolution for Semantic Image Segmentation. arXiv 2017.
Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation. arXiv 2018.

DeepLab V3, DeepLab V3+



출처 Rethinking Atrous Convolution for Semantic Image Segmentation. arXiv 2017.
Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation. arXiv 2018.

Segmentation Networks

- PASCAL VOC2012 Leaderboard

Model	Mean Average Precision(%)	Base CNN
DeepLab. V3+	87.8	Xception
DeepLab. V3	85.7	ResNet-101
PSPNet	85.4	ResNet-101
DeepLab. V2-CRF	79.7	ResNet-101
FCN-2s-Dilated_VGG19	69.0	VGG-19
FCN-8s	62.2	VGG-19
SegNet	59.9	VGG-19

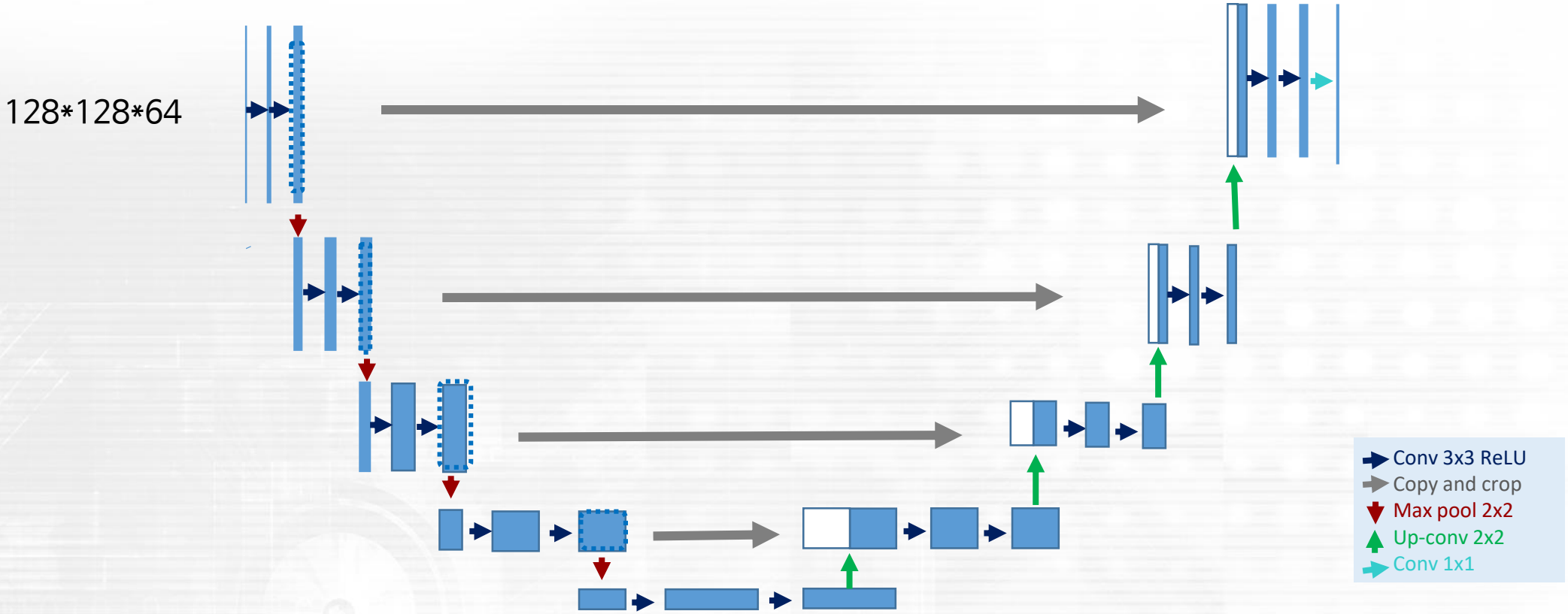
SPP

Dilated Conv

Encoder / Decoder

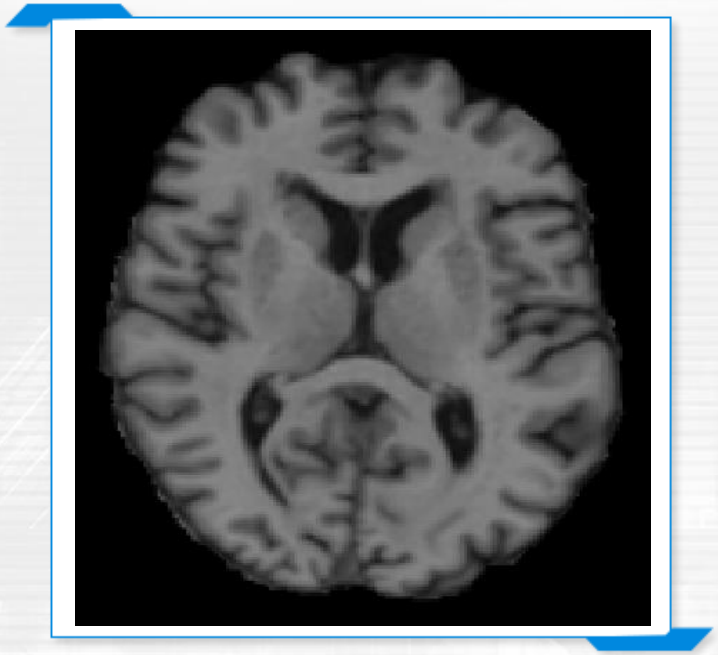
Encoder

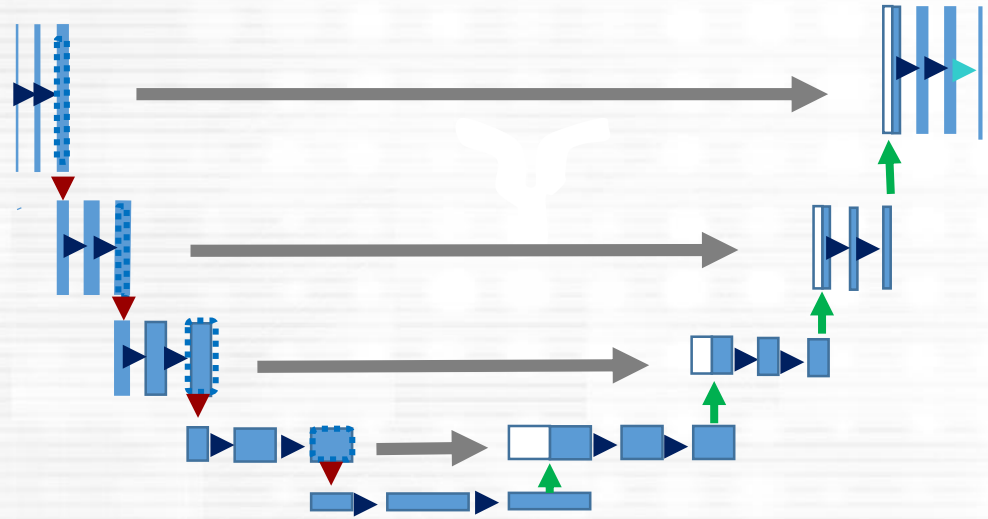
Segmentation in 3D



출처 V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation, arxiv 2016

Patch-wise segmentation





Loss Functions

$$CE(p,\hat{p}) = -(p\log(\hat{p}) + (1-p)\log(1-\hat{p}))$$

$$WCE(p,\hat{p}) = -(\beta p\log(\hat{p}) + (1-p)\log(1-\hat{p}))$$

$$BCE(p,\hat{p}) = -(\beta p\log(\hat{p}) + (1-\beta)(1-p)\log(1-\hat{p}))$$

Loss Functions

$$DL(p, \hat{p}) = 1 - \frac{2p\hat{p} + 1}{p + \hat{p} + 1}$$

$$DL(p, \hat{p}) = 1 - \frac{2\sum p_{h,w}\hat{p}_{h,w}}{\sum p_{h,w} + \sum \hat{p}_{h,w}}$$

Segmentation Metric

- Pixel accuracy

$$Accurary = \frac{TP + TN}{TP + TN + FP + FN}$$

- Intersection over Union

$$IoU = \frac{|S_g \cap S_r|}{|S_g \cup S_r|} = \frac{TP}{TP+FP+FN}$$

- Dice Coefficient

$$DICE = \frac{2|S_g \cap S_r|}{|S_g| + |S_r|} = \frac{2TP}{2TP+FP+FN}$$

Segmentation Metric

- Hausdorff distance

$$HD(A,B) = \max(h(A,B), h(B,A))$$

$$h(A,B) = \max_{a \in A} \min_{b \in B} \|a - b\|$$

$$AVD(A,B) = \max(d(A,B), d(B,A))$$

$$d(A,B) = \frac{1}{N} \sum_{a \in A} \min_{b \in B} \|a - b\|$$