

# Image Segmentation and Visualization

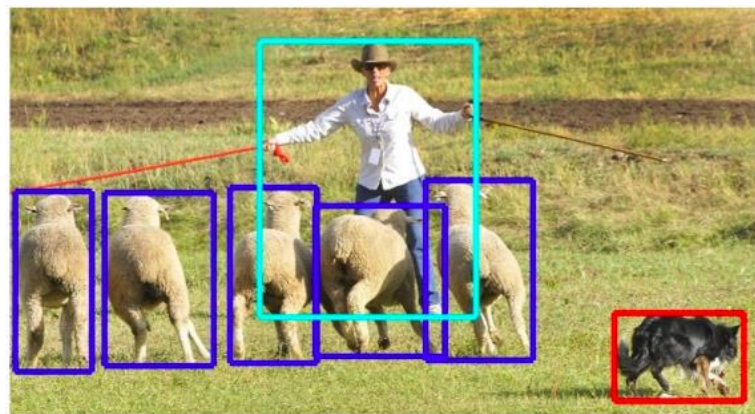
Xiaolong Wang

# Segmentation Problem and FCN

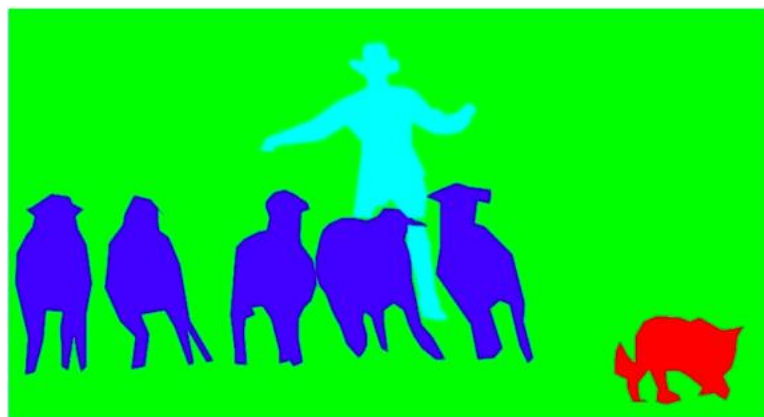
# The problem



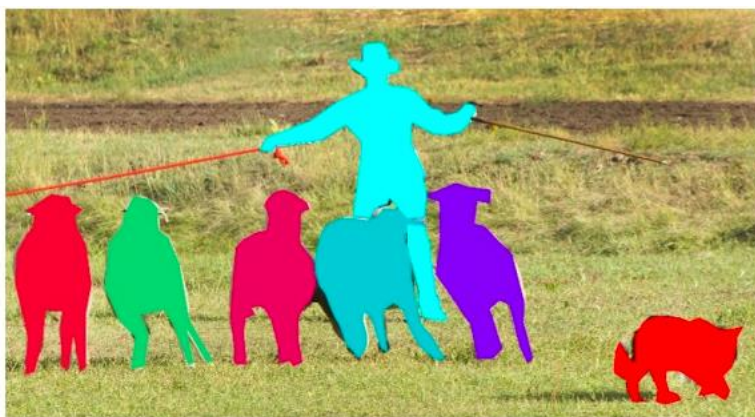
image classification



object detection

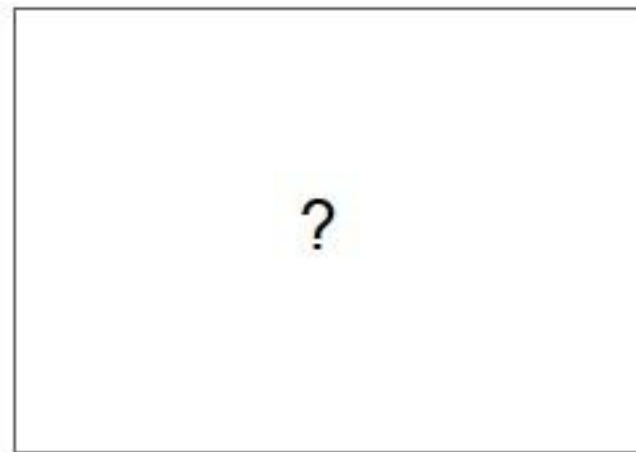
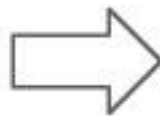


semantic segmentation



instance segmentation

# The problem



# Semantic Segmentation

Full image

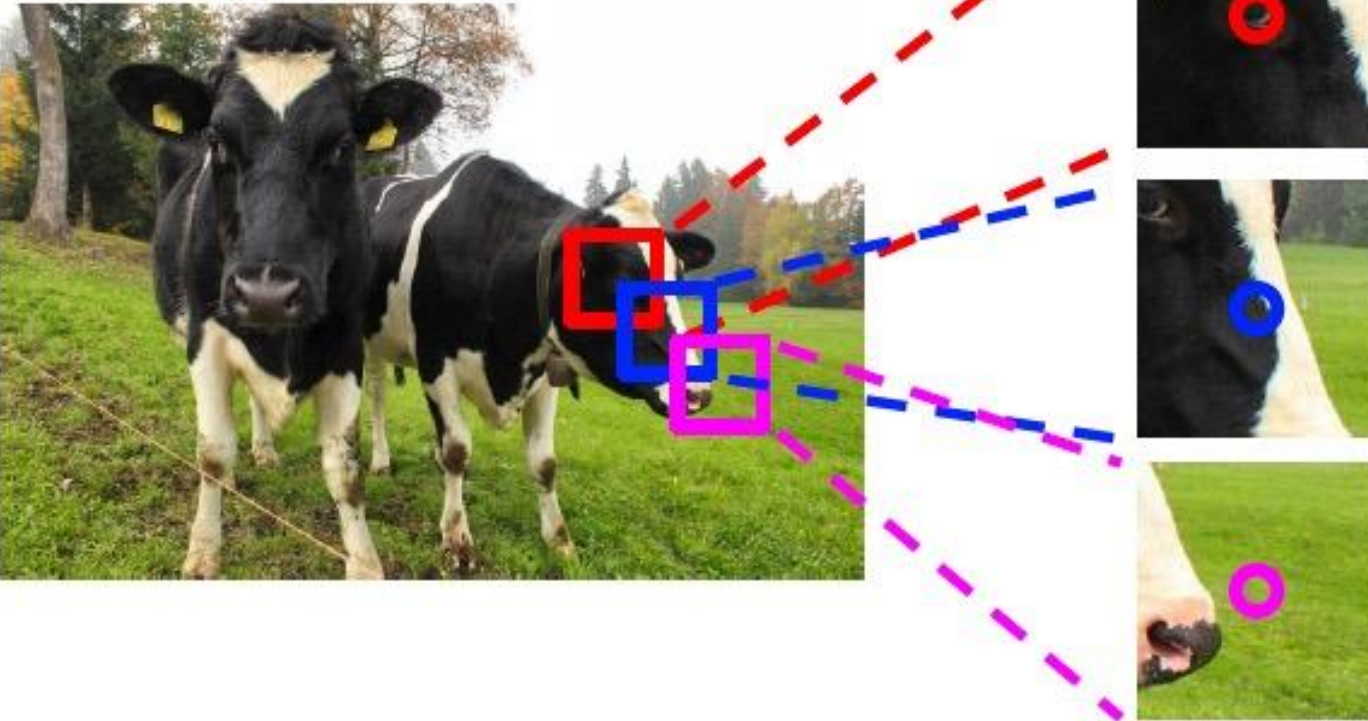


Simply staring one pixel is impossible to do the classification

Let's put in some context!

# Semantic Segmentation

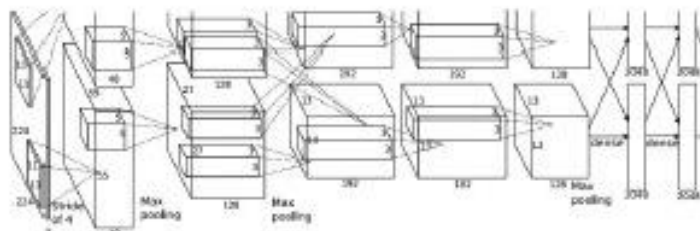
Full image



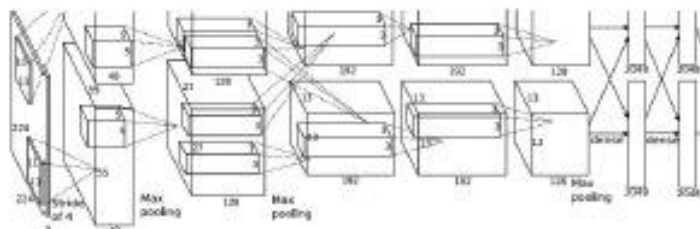


# Semantic Segmentation

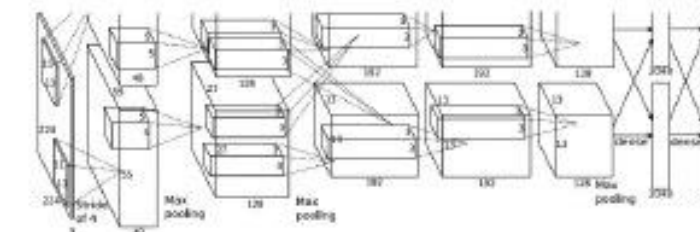
Full image



Cow



Cow

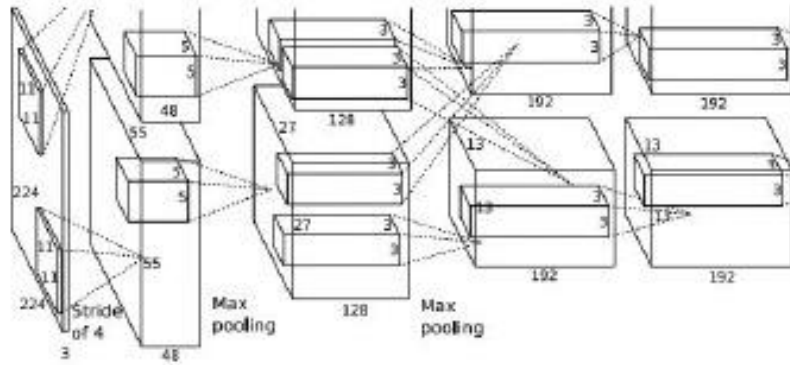


Grass

Time Consuming!

# Can we process the whole image at one time?

Full image



AlexNet input:  
 $227 \times 277 \times 3$

AlexNet Conv5:  
 $13 \times 13 \times 128$



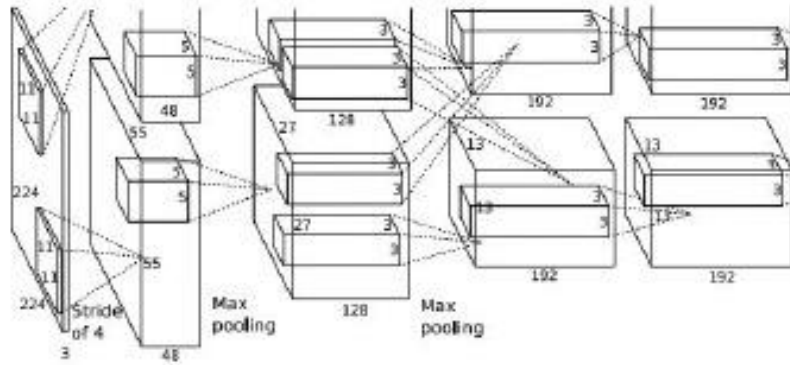
Output ?  
 $13 \times 13 \times 21$

**Output is too small!**



# Can we process the whole image at one time?

Full image



AlexNet input:  
 $227 \times 277 \times 3$

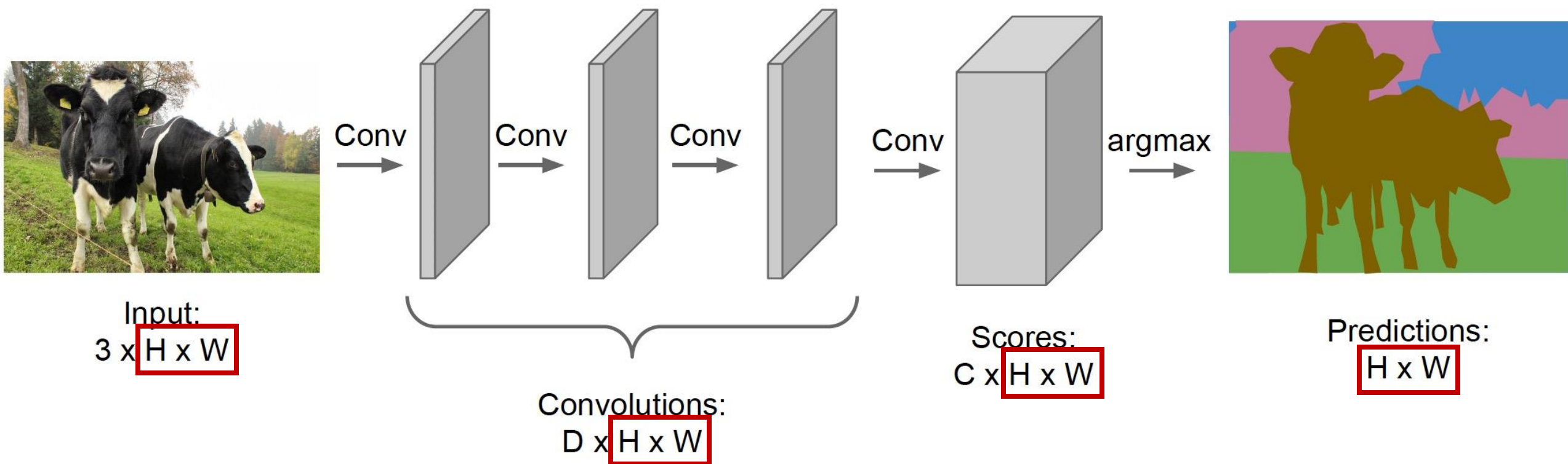
AlexNet Conv5:  
 $13 \times 13 \times 128$

FC +  
reshape  
→

Output ?  
 $227 \times 227 \times 21$

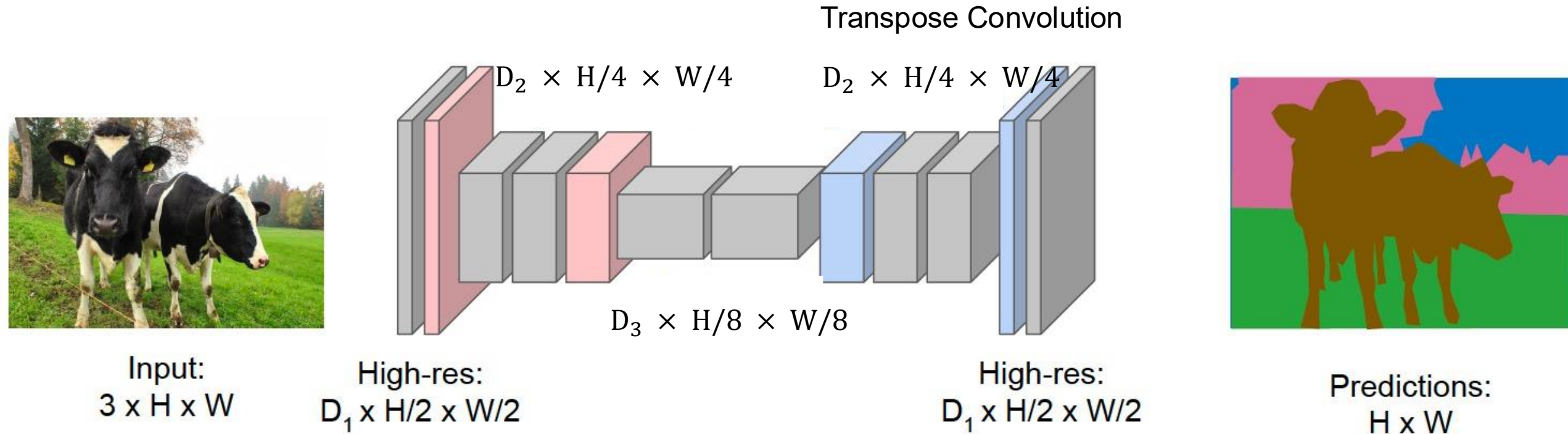
Huge number of Parameters for FC:  $13 \times 13 \times 128 \times 227 \times 227 \times 21$

# Fully Convolutional Network



Convolution at original image resolution has high computation cost.

# Fully Convolutional Network



Make the feature map small increases the receptive field

Make the feature map larger again increases the resolution

# The upsampling

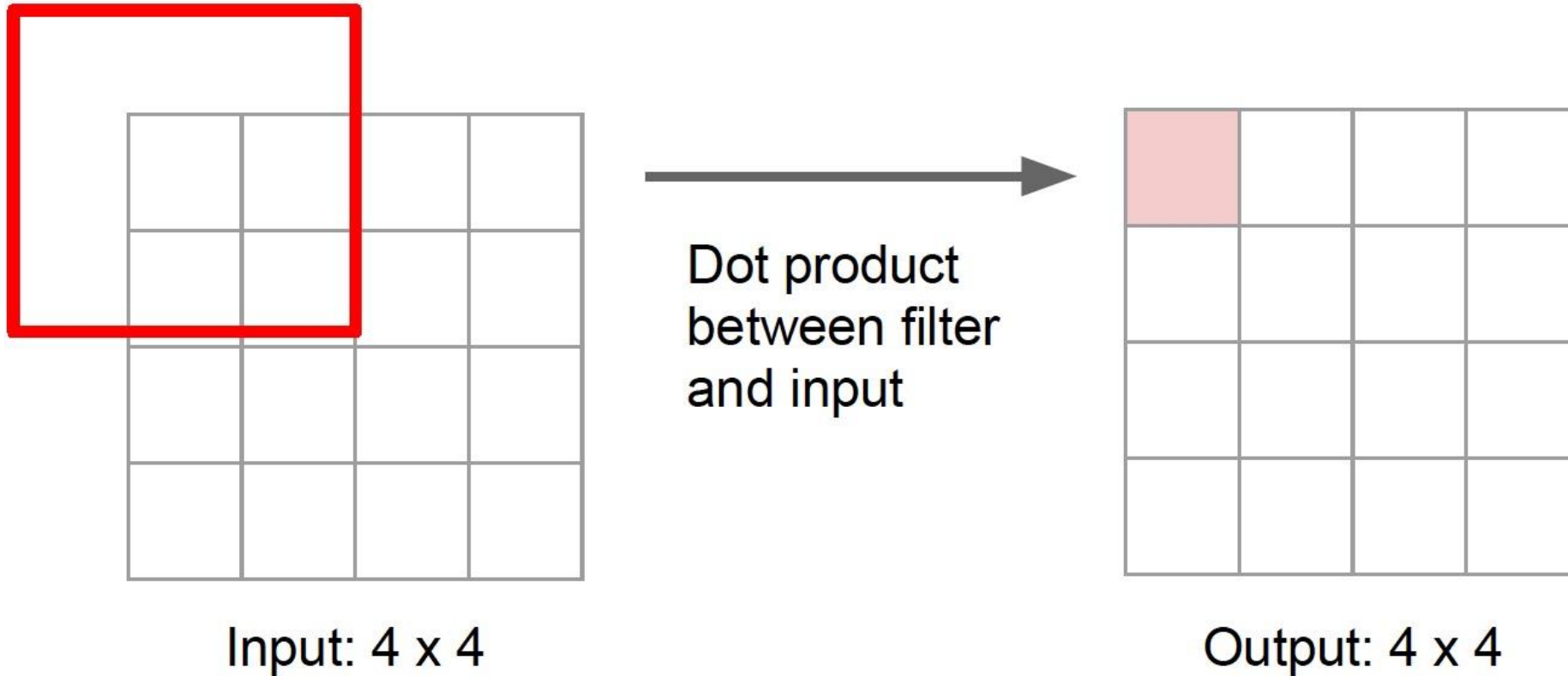
- Upsampling Layer
- Deconvolution Layer
- Transpose Convolution Layer

# Transpose Convolution



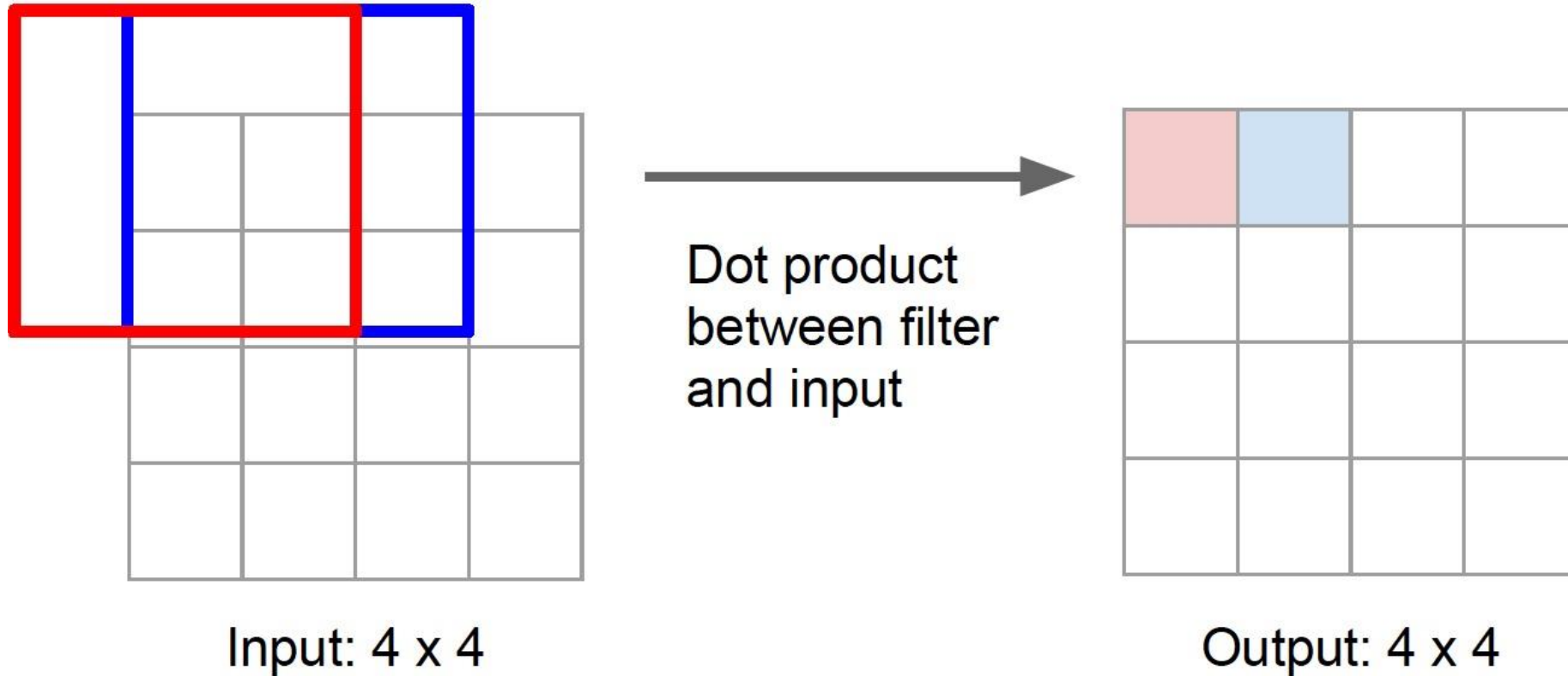
# Recall Convolution

3 X 3 convolution with stride 1 and padding 1



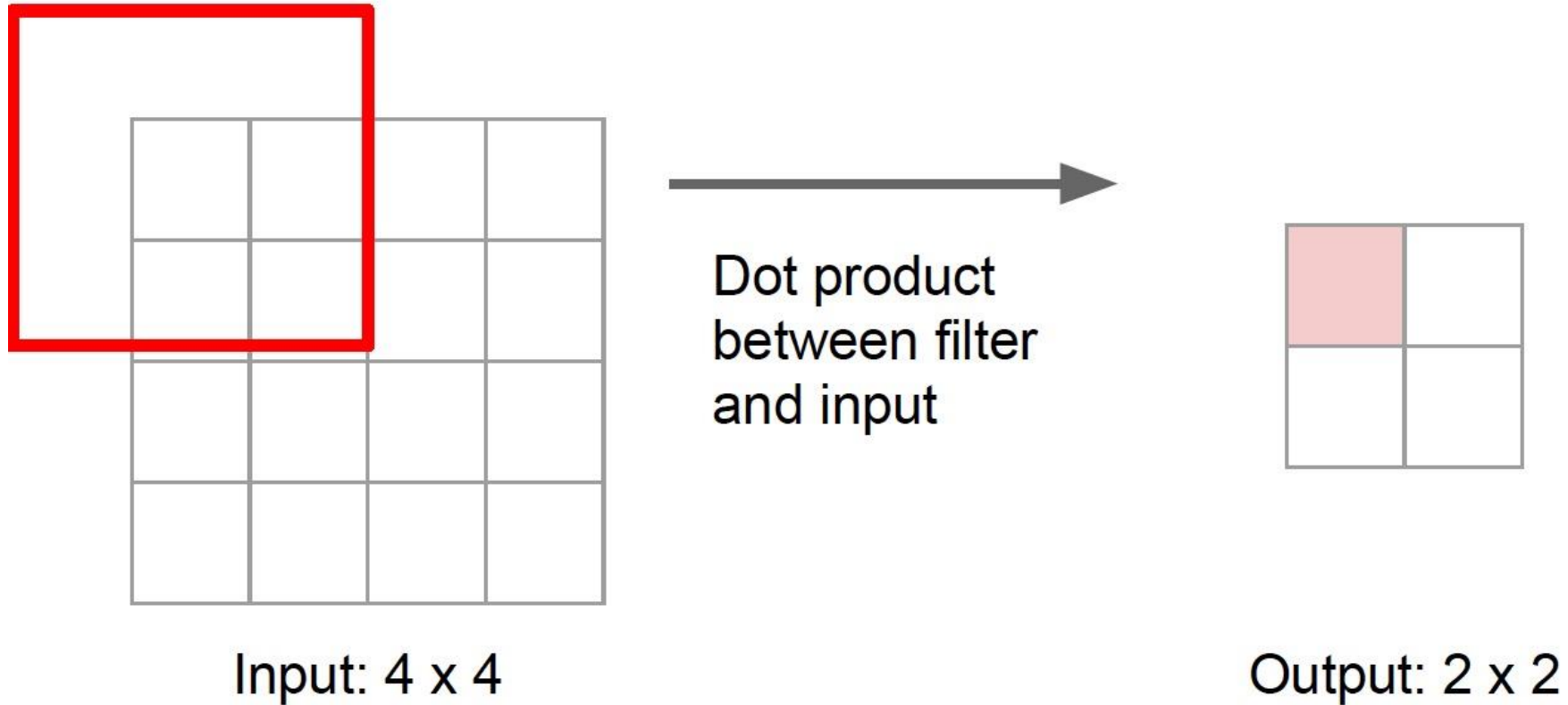
# Recall Convolution

3 X 3 convolution with stride 1 and padding 1



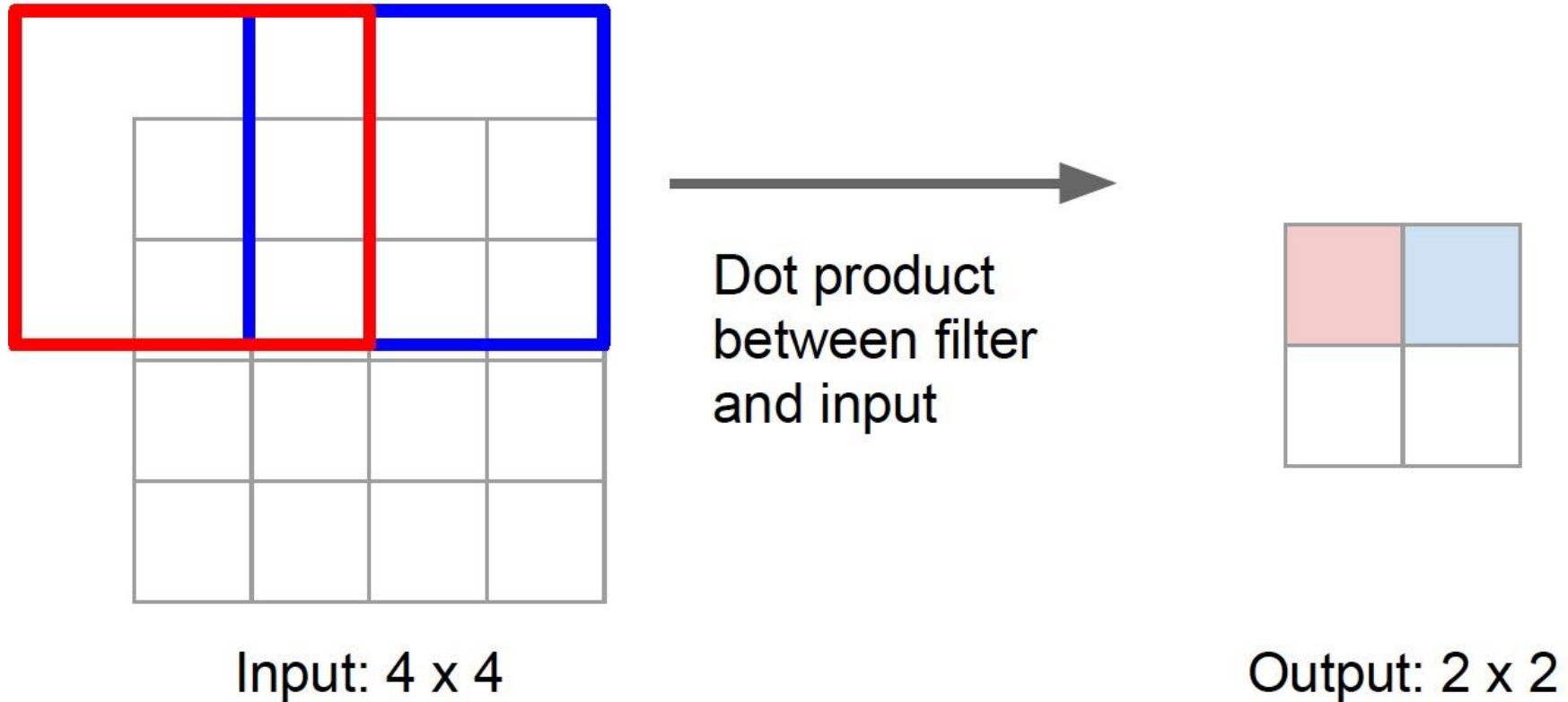
# Recall Convolution

3 X 3 convolution with stride 2 and padding 1



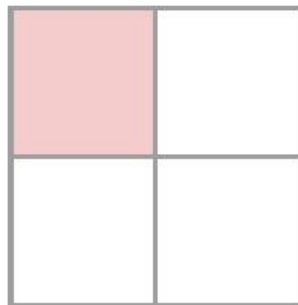
# Recall Convolution

3 X 3 convolution with stride 2 and padding 1



# Transpose Convolution

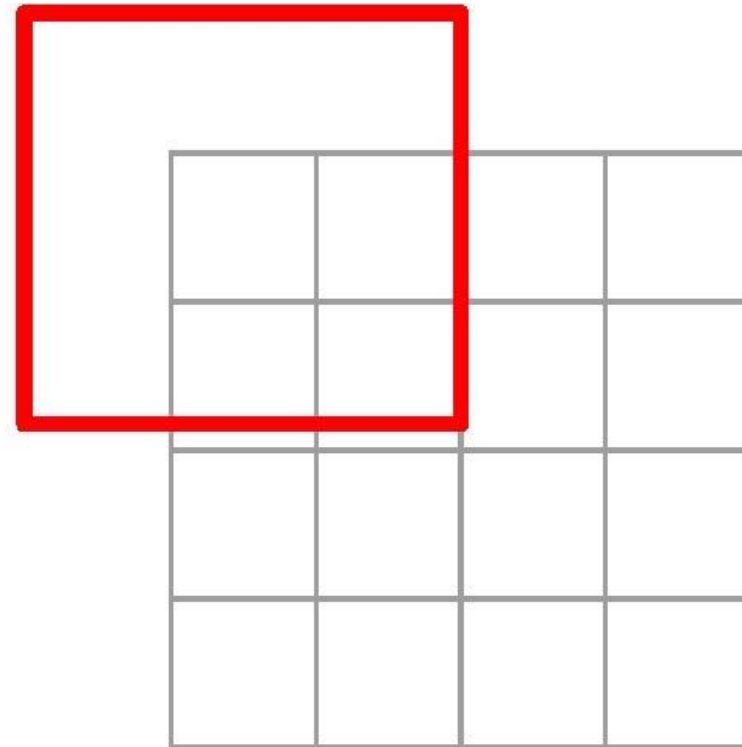
3 X 3 transpose convolution, stride 2 and padding 1



Input: 2 x 2



Input gives  
weight for  
filter

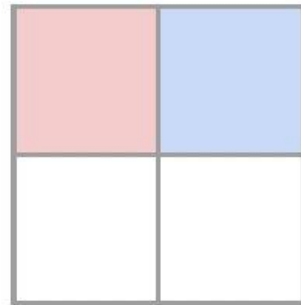


Output: 4 x 4



# Transpose Convolution

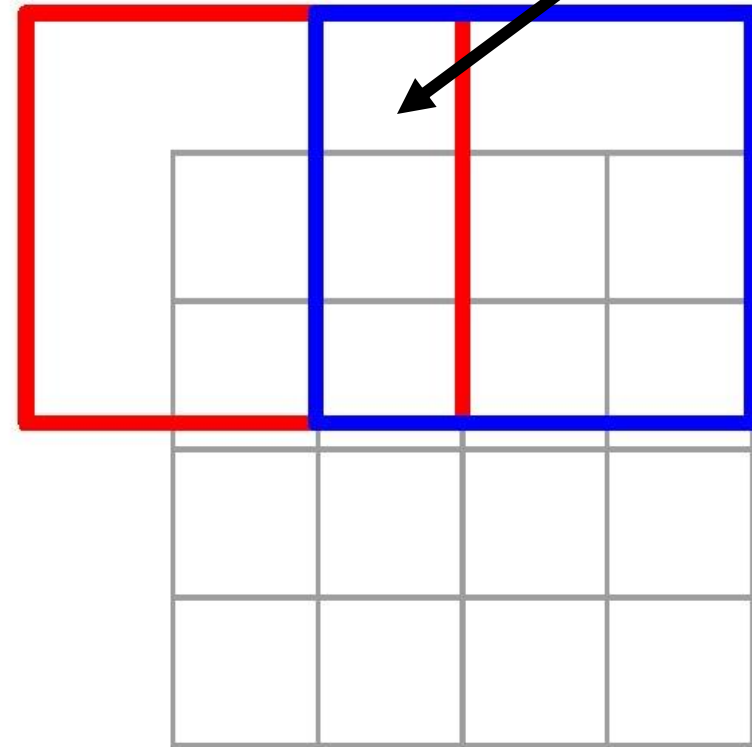
3 X 3 transpose convolution, stride 2 and padding 1



Input: 2 x 2



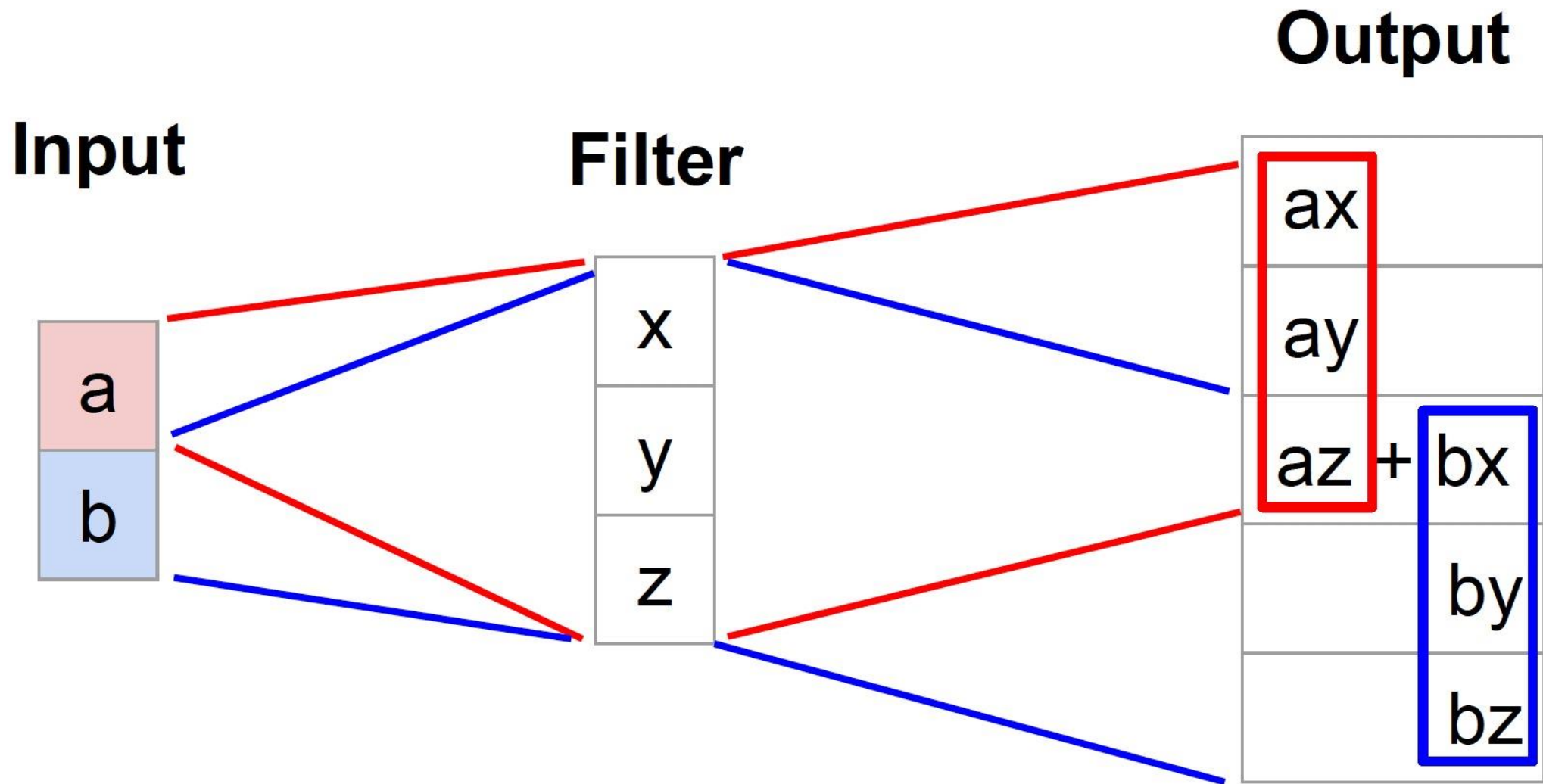
Input gives weight for filter



Output: 4 x 4

Sum over the overlapping region

# 1D Transpose Convolution Example



# 2D Convolution

Regular convolution (stride 1, pad 0)

$$\begin{array}{|c|c|c|c|} \hline x_{11} & x_{12} & x_{13} & x_{14} \\ \hline x_{21} & x_{22} & x_{23} & x_{24} \\ \hline x_{31} & x_{32} & x_{33} & x_{34} \\ \hline x_{41} & x_{42} & x_{43} & x_{44} \\ \hline \end{array}
 * \begin{array}{|c|c|c|} \hline w_{11} & w_{12} & w_{13} \\ \hline w_{21} & w_{22} & w_{23} \\ \hline w_{31} & w_{32} & w_{33} \\ \hline \end{array}
 = \begin{array}{|c|c|} \hline z_{11} & z_{12} \\ \hline z_{21} & z_{22} \\ \hline \end{array}$$

Matrix-vector form:

$$\begin{pmatrix} w_{11} & w_{12} & w_{13} & 0 & w_{21} & w_{22} & w_{23} & 0 & w_{31} & w_{32} & w_{33} & 0 & 0 & 0 & 0 & 0 \\ 0 & w_{11} & w_{12} & w_{13} & 0 & w_{21} & w_{22} & w_{23} & 0 & w_{31} & w_{32} & w_{33} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & w_{11} & w_{12} & w_{13} & 0 & w_{21} & w_{22} & w_{23} & 0 & w_{31} & w_{32} & w_{33} & 0 \\ 0 & 0 & 0 & 0 & 0 & w_{11} & w_{12} & w_{13} & 0 & w_{21} & w_{22} & w_{23} & 0 & w_{31} & w_{32} & w_{33} \end{pmatrix}
 \begin{pmatrix} x_{11} \\ x_{12} \\ x_{13} \\ x_{14} \\ \vdots \\ x_{44} \end{pmatrix}
 = \begin{pmatrix} z_{11} \\ z_{12} \\ z_{21} \\ z_{22} \end{pmatrix}$$

4x4 input, 2x2 output

# Transpose Convolution

[Source](#)

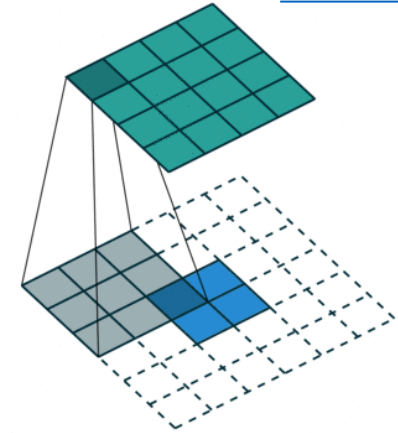
$z_{11}$	$z_{12}$
$z_{21}$	$z_{22}$

$*T$

$w_{11}$	$w_{12}$	$w_{13}$
$w_{21}$	$w_{22}$	$w_{23}$
$w_{31}$	$w_{32}$	$w_{33}$

=

$x_{11}$	$x_{12}$	$x_{13}$	$x_{14}$
$x_{21}$	$x_{22}$	$x_{23}$	$x_{24}$
$x_{31}$	$x_{32}$	$x_{33}$	$x_{34}$
$x_{41}$	$x_{42}$	$x_{43}$	$x_{44}$



$$\begin{pmatrix}
 w_{11} & 0 & 0 & 0 \\
 w_{12} & w_{11} & 0 & 0 \\
 w_{13} & w_{12} & 0 & 0 \\
 0 & w_{13} & 0 & 0 \\
 w_{21} & 0 & w_{11} & 0 \\
 w_{22} & w_{21} & w_{12} & w_{11} \\
 w_{23} & w_{22} & w_{13} & w_{12} \\
 0 & w_{23} & 0 & w_{13} \\
 w_{31} & 0 & w_{21} & 0 \\
 w_{32} & w_{31} & w_{22} & w_{21} \\
 w_{33} & w_{32} & w_{23} & w_{22} \\
 0 & w_{33} & 0 & w_{23} \\
 0 & 0 & w_{31} & 0 \\
 0 & 0 & w_{32} & w_{31} \\
 0 & 0 & w_{33} & w_{32} \\
 0 & 0 & 0 & w_{33}
 \end{pmatrix}
 \begin{pmatrix}
 z_{11} \\
 z_{12} \\
 z_{21} \\
 z_{22}
 \end{pmatrix}
 =
 \begin{pmatrix}
 x_{11} \\
 x_{12} \\
 x_{13} \\
 x_{14} \\
 x_{21} \\
 x_{22} \\
 x_{23} \\
 x_{24} \\
 x_{31} \\
 x_{32} \\
 x_{33} \\
 x_{34} \\
 x_{41} \\
 x_{42} \\
 x_{43} \\
 x_{44}
 \end{pmatrix}$$

2x2 input, 4x4 output

**Not** an inverse of the original convolution operation, simply reverses dimension change!

# Trans Convolution

$w_{33}$	$w_{32}$	$w_{31}$
$w_{23}$	$w_{22}$	$w_{21}$
$w_{13}$	$w_{12}$	$w_{11}$

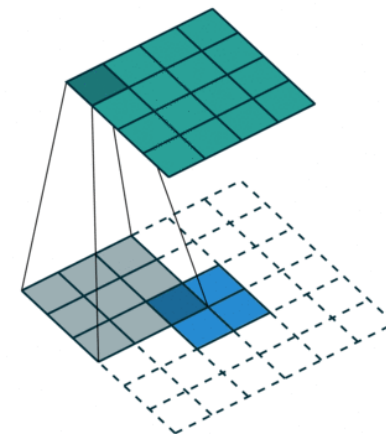
$z_{11}$	$z_{12}$
$z_{21}$	$z_{22}$

$*T$

$w_{11}$	$w_{12}$	$w_{13}$
$w_{21}$	$w_{22}$	$w_{23}$
$w_{31}$	$w_{32}$	$w_{33}$

=

$x_{11}$	$x_{12}$	$x_{13}$	$x_{14}$
$x_{21}$	$x_{22}$	$x_{23}$	$x_{24}$
$x_{31}$	$x_{32}$	$x_{33}$	$x_{34}$
$x_{41}$	$x_{42}$	$x_{43}$	$x_{44}$



$w_{11}$	0	0	0
$w_{12}$	$w_{11}$	0	0
$w_{13}$	$w_{12}$	0	0
0	$w_{13}$	0	0
$w_{21}$	0	$w_{11}$	0
$w_{22}$	$w_{21}$	$w_{12}$	$w_{11}$
$w_{23}$	$w_{22}$	$w_{13}$	$w_{12}$
0	$w_{23}$	0	$w_{13}$
$w_{31}$	0	$w_{21}$	0
$w_{32}$	$w_{31}$	$w_{22}$	$w_{21}$
$w_{33}$	$w_{32}$	$w_{23}$	$w_{22}$
0	$w_{33}$	0	$w_{23}$
0	0	$w_{31}$	0
0	0	$w_{32}$	$w_{31}$
0	0	$w_{33}$	$w_{32}$
0	0	0	$w_{33}$

$\begin{pmatrix} z_{11} \\ z_{12} \\ z_{21} \\ z_{22} \end{pmatrix}$

=

$x_{11}$
$x_{12}$
$x_{13}$
$x_{14}$
$x_{21}$
$x_{22}$
$x_{23}$
$x_{24}$
$x_{31}$
$x_{32}$
$x_{33}$
$x_{34}$
$x_{41}$
$x_{42}$
$x_{43}$
$x_{44}$

$x_{11} = w_{11}z_{11}$



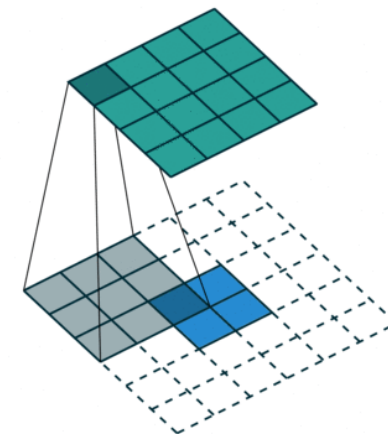
# Transposed Convolution

$w_{33}$	$w_{32}$	$w_{31}$
$w_{23}$	$w_{22}$	$w_{21}$
$w_{13}$	$w_{12}$	$w_{11}$

$z_{11}$	$z_{12}$
$z_{21}$	$z_{22}$

=

$x_{11}$	$x_{12}$	$x_{13}$	$x_{14}$
$x_{21}$	$x_{22}$	$x_{23}$	$x_{24}$
$x_{31}$	$x_{32}$	$x_{33}$	$x_{34}$
$x_{41}$	$x_{42}$	$x_{43}$	$x_{44}$



$w_{11}$	0	0	0
$w_{12}$	$w_{11}$	0	0
$w_{13}$	$w_{12}$	0	0
0	$w_{13}$	0	0
$w_{21}$	0	$w_{11}$	0
$w_{22}$	$w_{21}$	$w_{12}$	$w_{11}$
$w_{23}$	$w_{22}$	$w_{13}$	$w_{12}$
0	$w_{23}$	0	$w_{13}$
$w_{31}$	0	$w_{21}$	0
$w_{32}$	$w_{31}$	$w_{22}$	$w_{21}$
$w_{33}$	$w_{32}$	$w_{23}$	$w_{22}$
0	$w_{33}$	0	$w_{23}$
0	0	$w_{31}$	0
0	0	$w_{32}$	$w_{31}$
0	0	$w_{33}$	$w_{32}$
0	0	0	$w_{33}$

$z_{11}$
$z_{12}$
$z_{21}$
$z_{22}$

=

$x_{11}$
$x_{12}$
$x_{13}$
$x_{14}$
$x_{21}$
$x_{22}$
$x_{23}$
$x_{24}$
$x_{31}$
$x_{32}$
$x_{33}$
$x_{34}$
$x_{41}$
$x_{42}$
$x_{43}$
$x_{44}$

$$x_{11} = w_{11}z_{11}$$

# Transposed convolution

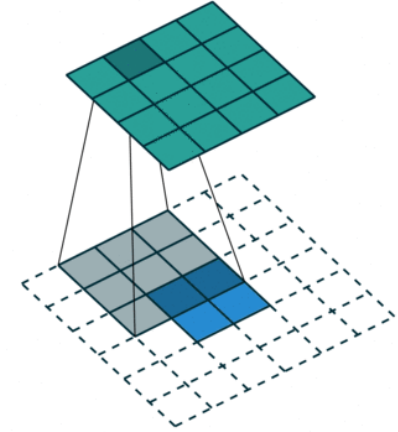
$w_{33}$	$w_{32}$	$w_{31}$
$w_{23}$	$w_{22}$	$w_{21}$
$w_{13}$	$w_{12}$	$w_{11}$
$z_{11}$	$z_{12}$	
$z_{21}$	$z_{22}$	

$*T$

$w_{11}$	$w_{12}$	$w_{13}$
$w_{21}$	$w_{22}$	$w_{23}$
$w_{31}$	$w_{32}$	$w_{33}$

=

$x_{11}$	$x_{12}$	$x_{13}$	$x_{14}$
$x_{21}$	$x_{22}$	$x_{23}$	$x_{24}$
$x_{31}$	$x_{32}$	$x_{33}$	$x_{34}$
$x_{41}$	$x_{42}$	$x_{43}$	$x_{44}$

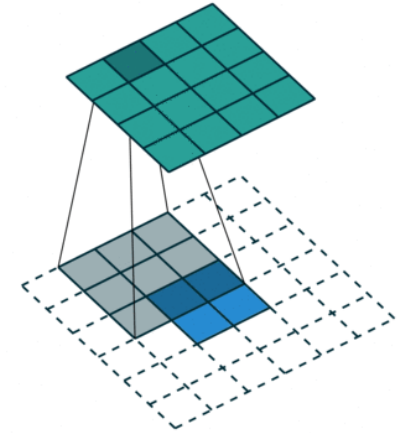
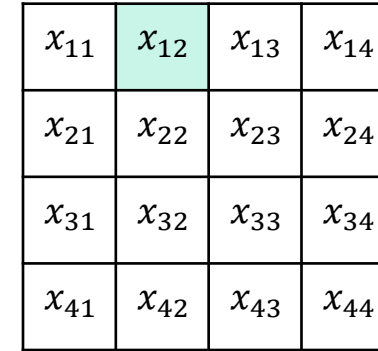
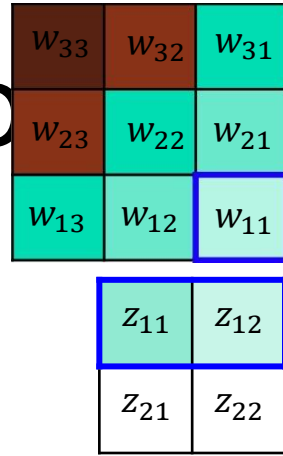


Convolve input with *flipped* filter

$$x_{12} = w_{12}z_{11} + w_{11}z_{12}$$

$w_{11}$	0	0	0	$\left( \begin{matrix} z_{11} \\ z_{12} \\ z_{21} \\ z_{22} \end{matrix} \right) =$	$x_{11}$
$w_{12}$	$w_{11}$	0	0		$x_{12}$
$w_{13}$	$w_{12}$	0	0		$x_{13}$
0	$w_{13}$	0	0		$x_{14}$
$w_{21}$	0	$w_{11}$	0		$x_{21}$
$w_{22}$	$w_{21}$	$w_{12}$	$w_{11}$		$x_{22}$
$w_{23}$	$w_{22}$	$w_{13}$	$w_{12}$		$x_{23}$
0	$w_{23}$	0	$w_{13}$		$x_{24}$
$w_{31}$	0	$w_{21}$	0		$x_{31}$
$w_{32}$	$w_{31}$	$w_{22}$	$w_{21}$		$x_{32}$
$w_{33}$	$w_{32}$	$w_{23}$	$w_{22}$		$x_{33}$
0	$w_{33}$	0	$w_{23}$		$x_{34}$
0	0	$w_{31}$	0		$x_{41}$
0	0	$w_{32}$	$w_{31}$		$x_{42}$
0	0	$w_{33}$	$w_{32}$	$x_{43}$	
0	0	0	$w_{33}$	$x_{44}$	

# Transposed Convolution

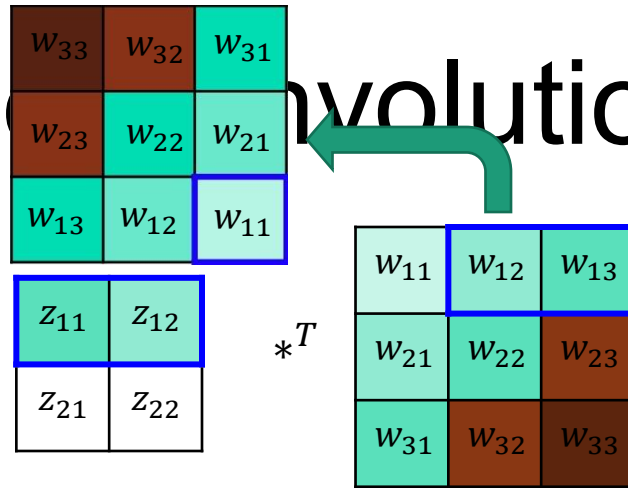


Convolve input with *flipped* filter

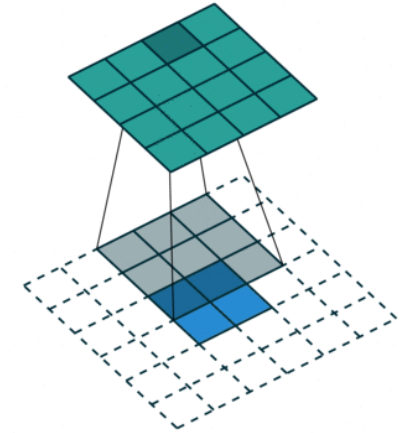
$$x_{12} = w_{12}z_{11} + w_{11}z_{12}$$

$w_{11}$	0	0	0	$\left( \begin{array}{c} z_{11} \\ z_{12} \\ z_{21} \\ z_{22} \end{array} \right)$	$=$	$x_{11}$
$w_{12}$	$w_{11}$	0	0			$x_{12}$
$w_{13}$	$w_{12}$	0	0			$x_{13}$
0	$w_{13}$	0	0			$x_{14}$
$w_{21}$	0	$w_{11}$	0			$x_{21}$
$w_{22}$	$w_{21}$	$w_{12}$	$w_{11}$			$x_{22}$
$w_{23}$	$w_{22}$	$w_{13}$	$w_{12}$			$x_{23}$
0	$w_{23}$	0	$w_{13}$			$x_{24}$
$w_{31}$	0	$w_{21}$	0			$x_{31}$
$w_{32}$	$w_{31}$	$w_{22}$	$w_{21}$			$x_{32}$
$w_{33}$	$w_{32}$	$w_{23}$	$w_{22}$			$x_{33}$
0	$w_{33}$	0	$w_{23}$			$x_{34}$
0	0	$w_{31}$	0			$x_{41}$
0	0	$w_{32}$	$w_{31}$			$x_{42}$
0	0	$w_{33}$	$w_{32}$			$x_{43}$
0	0	0	$w_{33}$			$x_{44}$

# Transposition Convolution



$x_{11}$	$x_{12}$	$x_{13}$	$x_{14}$
$x_{21}$	$x_{22}$	$x_{23}$	$x_{24}$
$x_{31}$	$x_{32}$	$x_{33}$	$x_{34}$
$x_{41}$	$x_{42}$	$x_{43}$	$x_{44}$



Convolve input with *flipped* filter

$$\begin{pmatrix}
 w_{11} & 0 & 0 & 0 \\
 w_{12} & w_{11} & 0 & 0 \\
 w_{13} & w_{12} & 0 & 0 \\
 0 & w_{13} & 0 & 0 \\
 w_{21} & 0 & w_{11} & 0 \\
 w_{22} & w_{21} & w_{12} & w_{11} \\
 w_{23} & w_{22} & w_{13} & w_{12} \\
 0 & w_{23} & 0 & w_{13} \\
 w_{31} & 0 & w_{21} & 0 \\
 w_{32} & w_{31} & w_{22} & w_{21} \\
 w_{33} & w_{32} & w_{23} & w_{22} \\
 0 & w_{33} & 0 & w_{23} \\
 0 & 0 & w_{31} & 0 \\
 0 & 0 & w_{32} & w_{31} \\
 0 & 0 & w_{33} & w_{32} \\
 0 & 0 & 0 & w_{33}
 \end{pmatrix}
 \begin{pmatrix}
 z_{11} \\
 z_{12} \\
 z_{21} \\
 z_{22}
 \end{pmatrix}
 =
 \begin{pmatrix}
 x_{11} \\
 x_{12} \\
 x_{13} \\
 x_{14} \\
 x_{21} \\
 x_{22} \\
 x_{23} \\
 x_{24} \\
 x_{31} \\
 x_{32} \\
 x_{33} \\
 x_{34} \\
 x_{41} \\
 x_{42} \\
 x_{43} \\
 x_{44}
 \end{pmatrix}$$

$$x_{13} = w_{13}z_{11} + w_{12}z_{12}$$

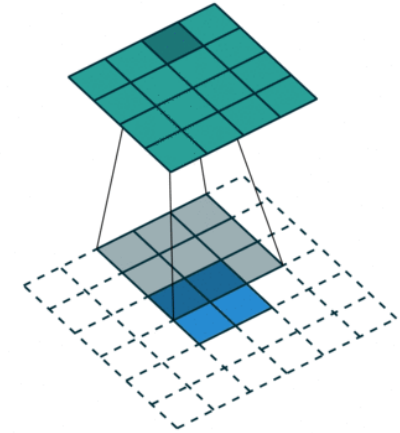
# Transposed Convolution

$w_{33}$	$w_{32}$	$w_{31}$
$w_{23}$	$w_{22}$	$w_{21}$
$w_{13}$	$w_{12}$	$w_{11}$
$z_{11}$	$z_{12}$	
$z_{21}$	$z_{22}$	

$x_{11}$	$x_{12}$	$x_{13}$	$x_{14}$
$x_{21}$	$x_{22}$	$x_{23}$	$x_{24}$
$x_{31}$	$x_{32}$	$x_{33}$	$x_{34}$
$x_{41}$	$x_{42}$	$x_{43}$	$x_{44}$

=

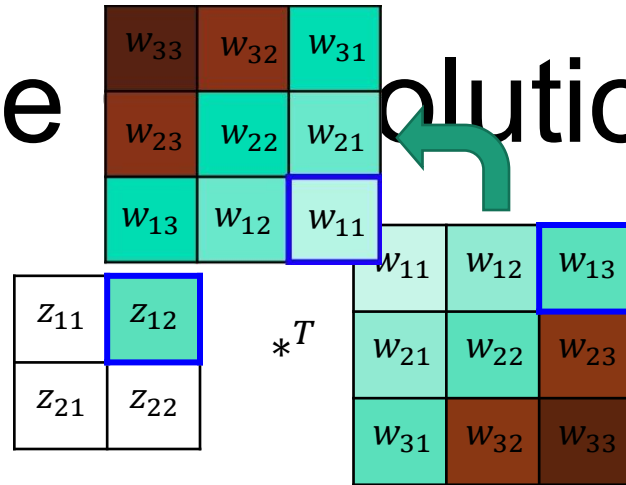
Convolve input with *flipped* filter



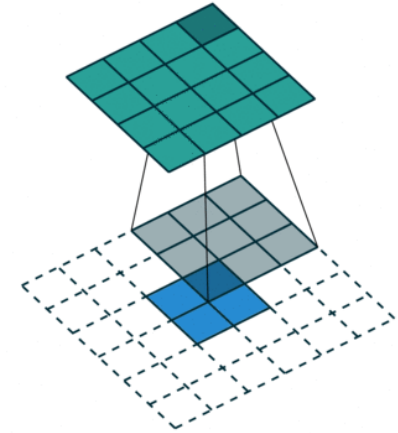
$$\begin{pmatrix}
 w_{11} & 0 & 0 & 0 \\
 w_{12} & w_{11} & 0 & 0 \\
 w_{13} & w_{12} & 0 & 0 \\
 0 & w_{13} & 0 & 0 \\
 w_{21} & 0 & w_{11} & 0 \\
 w_{22} & w_{21} & w_{12} & w_{11} \\
 w_{23} & w_{22} & w_{13} & w_{12} \\
 0 & w_{23} & 0 & w_{13} \\
 w_{31} & 0 & w_{21} & 0 \\
 w_{32} & w_{31} & w_{22} & w_{21} \\
 w_{33} & w_{32} & w_{23} & w_{22} \\
 0 & w_{33} & 0 & w_{23} \\
 0 & 0 & w_{31} & 0 \\
 0 & 0 & w_{32} & w_{31} \\
 0 & 0 & w_{33} & w_{32} \\
 0 & 0 & 0 & w_{33}
 \end{pmatrix}
 \begin{pmatrix}
 z_{11} \\
 z_{12} \\
 z_{21} \\
 z_{22}
 \end{pmatrix}
 =
 \begin{pmatrix}
 x_{11} \\
 x_{12} \\
 x_{13} \\
 x_{14} \\
 x_{21} \\
 x_{22} \\
 x_{23} \\
 x_{24} \\
 x_{31} \\
 x_{32} \\
 x_{33} \\
 x_{34} \\
 x_{41} \\
 x_{42} \\
 x_{43} \\
 x_{44}
 \end{pmatrix}$$

$$x_{13} = w_{13}z_{11} + w_{12}z_{12}$$

# Transpose Convolution



$x_{11}$	$x_{12}$	$x_{13}$	$x_{14}$
$x_{21}$	$x_{22}$	$x_{23}$	$x_{24}$
$x_{31}$	$x_{32}$	$x_{33}$	$x_{34}$
$x_{41}$	$x_{42}$	$x_{43}$	$x_{44}$



Convolve input with *flipped* filter

$$\begin{pmatrix}
 w_{11} & 0 & 0 & 0 \\
 w_{12} & w_{11} & 0 & 0 \\
 w_{13} & w_{12} & 0 & 0 \\
 \mathbf{0} & \mathbf{w_{13}} & 0 & 0 \\
 w_{21} & 0 & w_{11} & 0 \\
 w_{22} & w_{21} & w_{12} & w_{11} \\
 w_{23} & w_{22} & w_{13} & w_{12} \\
 0 & w_{23} & 0 & w_{13} \\
 w_{31} & 0 & w_{21} & 0 \\
 w_{32} & w_{31} & w_{22} & w_{21} \\
 w_{33} & w_{32} & w_{23} & w_{22} \\
 0 & w_{33} & 0 & w_{23} \\
 0 & 0 & w_{31} & 0 \\
 0 & 0 & w_{32} & w_{31} \\
 0 & 0 & w_{33} & w_{32} \\
 0 & 0 & 0 & w_{33}
 \end{pmatrix}
 \begin{pmatrix}
 z_{11} \\
 z_{12} \\
 z_{21} \\
 z_{22}
 \end{pmatrix}
 =
 \begin{pmatrix}
 x_{11} \\
 x_{12} \\
 x_{13} \\
 \mathbf{x_{14}} \\
 x_{21} \\
 x_{22} \\
 x_{23} \\
 x_{24} \\
 x_{31} \\
 x_{32} \\
 x_{33} \\
 x_{34} \\
 x_{41} \\
 x_{42} \\
 x_{43} \\
 x_{44}
 \end{pmatrix}$$

$x_{14} = w_{13}z_{12}$

# Trans Convolution

$w_{33}$	$w_{32}$	$w_{31}$
$w_{23}$	$w_{22}$	$w_{21}$
$w_{13}$	$w_{12}$	$w_{11}$

$z_{11}$	$z_{12}$
$z_{21}$	$z_{22}$

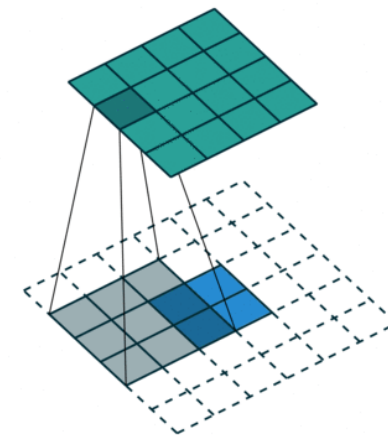
$*^T$

$w_{11}$	$w_{12}$	$w_{13}$
$w_{21}$	$w_{22}$	$w_{23}$
$w_{31}$	$w_{32}$	$w_{33}$

=

$x_{11}$	$x_{12}$	$x_{13}$	$x_{14}$
$x_{21}$	$x_{22}$	$x_{23}$	$x_{24}$
$x_{31}$	$x_{32}$	$x_{33}$	$x_{34}$
$x_{41}$	$x_{42}$	$x_{43}$	$x_{44}$

Convolve input with *flipped* filter

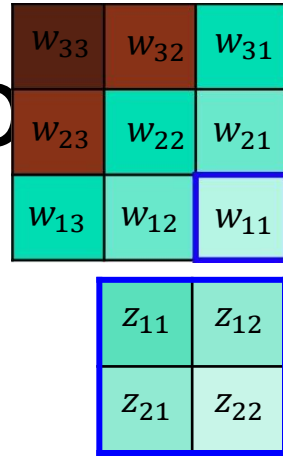


$$\begin{pmatrix}
 w_{11} & 0 & 0 & 0 \\
 w_{12} & w_{11} & 0 & 0 \\
 w_{13} & w_{12} & 0 & 0 \\
 0 & w_{13} & 0 & 0 \\
 w_{21} & 0 & w_{11} & 0 \\
 w_{22} & w_{21} & w_{12} & w_{11} \\
 w_{23} & w_{22} & w_{13} & w_{12} \\
 0 & w_{23} & 0 & w_{13} \\
 w_{31} & 0 & w_{21} & 0 \\
 w_{32} & w_{31} & w_{22} & w_{21} \\
 w_{33} & w_{32} & w_{23} & w_{22} \\
 0 & w_{33} & 0 & w_{23} \\
 0 & 0 & w_{31} & 0 \\
 0 & 0 & w_{32} & w_{31} \\
 0 & 0 & w_{33} & w_{32} \\
 0 & 0 & 0 & w_{33}
 \end{pmatrix}
 \begin{pmatrix}
 z_{11} \\
 z_{12} \\
 z_{21} \\
 z_{22}
 \end{pmatrix}
 =
 \begin{pmatrix}
 x_{11} \\
 x_{12} \\
 x_{13} \\
 x_{14} \\
 x_{21} \\
 x_{22} \\
 x_{23} \\
 x_{24} \\
 x_{31} \\
 x_{32} \\
 x_{33} \\
 x_{34} \\
 x_{41} \\
 x_{42} \\
 x_{43} \\
 x_{44}
 \end{pmatrix}$$

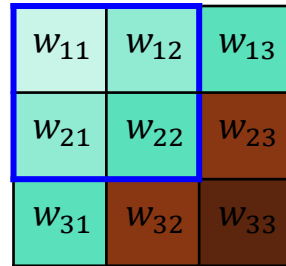
$$x_{21} = w_{21}z_{11} + w_{11}z_{21}$$



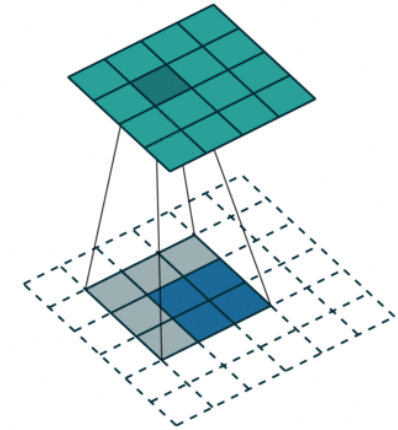
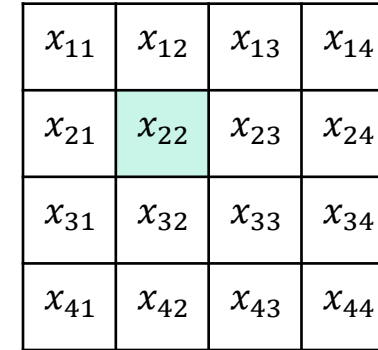
# Transposed convolution



$*T$



=



Convolve input with *flipped* filter

$w_{11}$	0	0	0
$w_{12}$	$w_{11}$	0	0
$w_{13}$	$w_{12}$	0	0
0	$w_{13}$	0	0
$w_{21}$	0	$w_{11}$	0
$w_{22}$	$w_{21}$	$w_{12}$	$w_{11}$
$w_{23}$	$w_{22}$	$w_{13}$	$w_{12}$
0	$w_{23}$	0	$w_{13}$
$w_{31}$	0	$w_{21}$	0
$w_{32}$	$w_{31}$	$w_{22}$	$w_{21}$
$w_{33}$	$w_{32}$	$w_{23}$	$w_{22}$
0	$w_{33}$	0	$w_{23}$
0	0	$w_{31}$	0
0	0	$w_{32}$	$w_{31}$
0	0	$w_{33}$	$w_{32}$
0	0	0	$w_{33}$

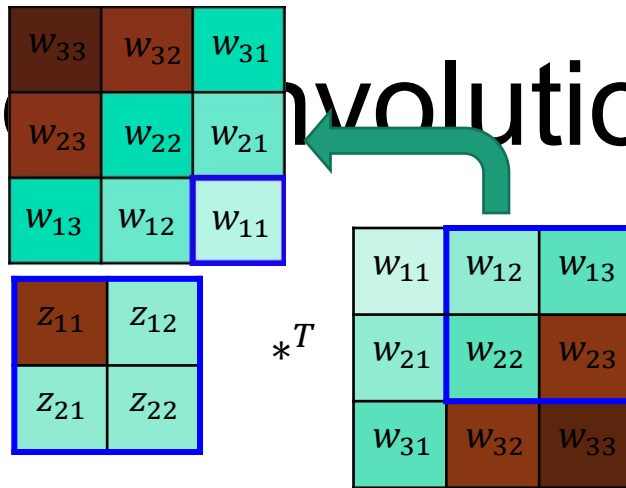
$\left. \begin{matrix} z_{11} \\ z_{12} \\ z_{21} \\ z_{22} \end{matrix} \right\}$

=

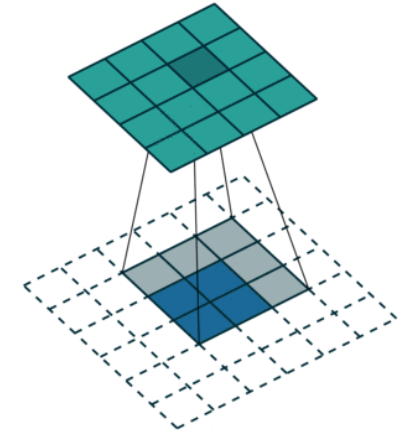
$x_{11}$
$x_{12}$
$x_{13}$
$x_{14}$
$x_{21}$
$x_{22}$
$x_{23}$
$x_{24}$
$x_{31}$
$x_{32}$
$x_{33}$
$x_{34}$
$x_{41}$
$x_{42}$
$x_{43}$
$x_{44}$

$$x_{22} = w_{22}z_{11} + w_{21}z_{12} + w_{12}z_{21} + w_{11}z_{22}$$

# Transposed Convolution



$x_{11}$	$x_{12}$	$x_{13}$	$x_{14}$
$x_{21}$	$x_{22}$	$x_{23}$	$x_{24}$
$x_{31}$	$x_{32}$	$x_{33}$	$x_{34}$
$x_{41}$	$x_{42}$	$x_{43}$	$x_{44}$

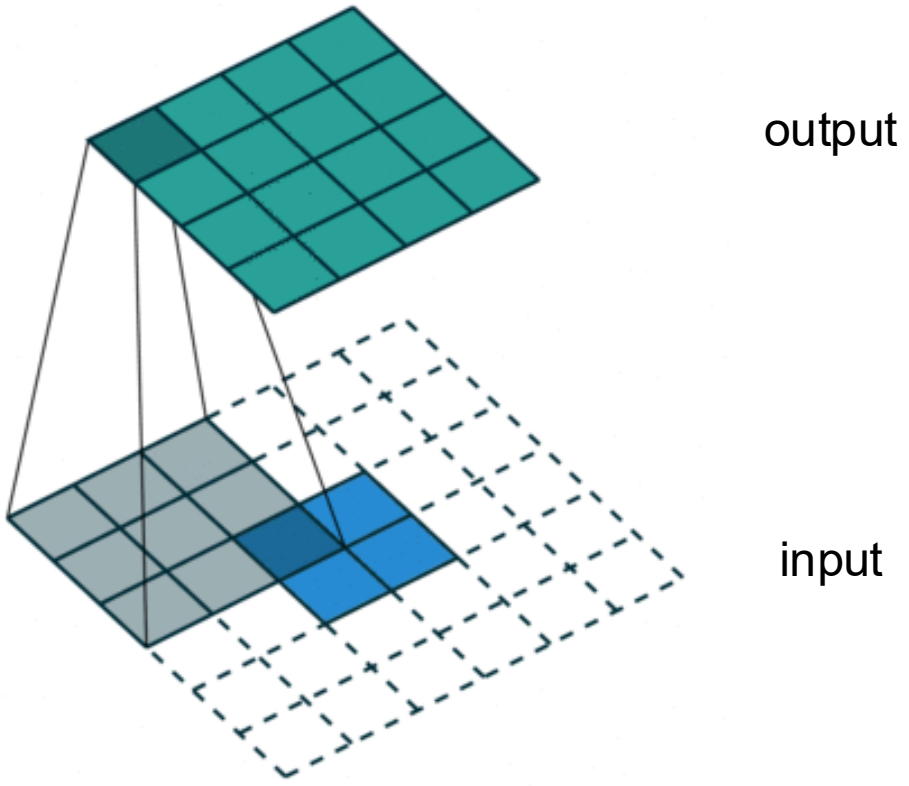


Convolve input with *flipped* filter

$$\begin{pmatrix}
 w_{11} & 0 & 0 & 0 \\
 w_{12} & w_{11} & 0 & 0 \\
 w_{13} & w_{12} & 0 & 0 \\
 0 & w_{13} & 0 & 0 \\
 w_{21} & 0 & w_{11} & 0 \\
 w_{22} & w_{21} & w_{12} & w_{11} \\
 w_{23} & w_{22} & w_{13} & w_{12} \\
 0 & w_{23} & 0 & w_{13} \\
 w_{31} & 0 & w_{21} & 0 \\
 w_{32} & w_{31} & w_{22} & w_{21} \\
 w_{33} & w_{32} & w_{23} & w_{22} \\
 0 & w_{33} & 0 & w_{23} \\
 0 & 0 & w_{31} & 0 \\
 0 & 0 & w_{32} & w_{31} \\
 0 & 0 & w_{33} & w_{32} \\
 0 & 0 & 0 & w_{33}
 \end{pmatrix}
 \begin{pmatrix}
 z_{11} \\
 z_{12} \\
 z_{21} \\
 z_{22}
 \end{pmatrix}
 =
 \begin{pmatrix}
 x_{11} \\
 x_{12} \\
 x_{13} \\
 x_{14} \\
 x_{21} \\
 x_{22} \\
 x_{23} \\
 x_{24} \\
 x_{31} \\
 x_{32} \\
 x_{33} \\
 x_{34} \\
 x_{41} \\
 x_{42} \\
 x_{43} \\
 x_{44}
 \end{pmatrix}$$

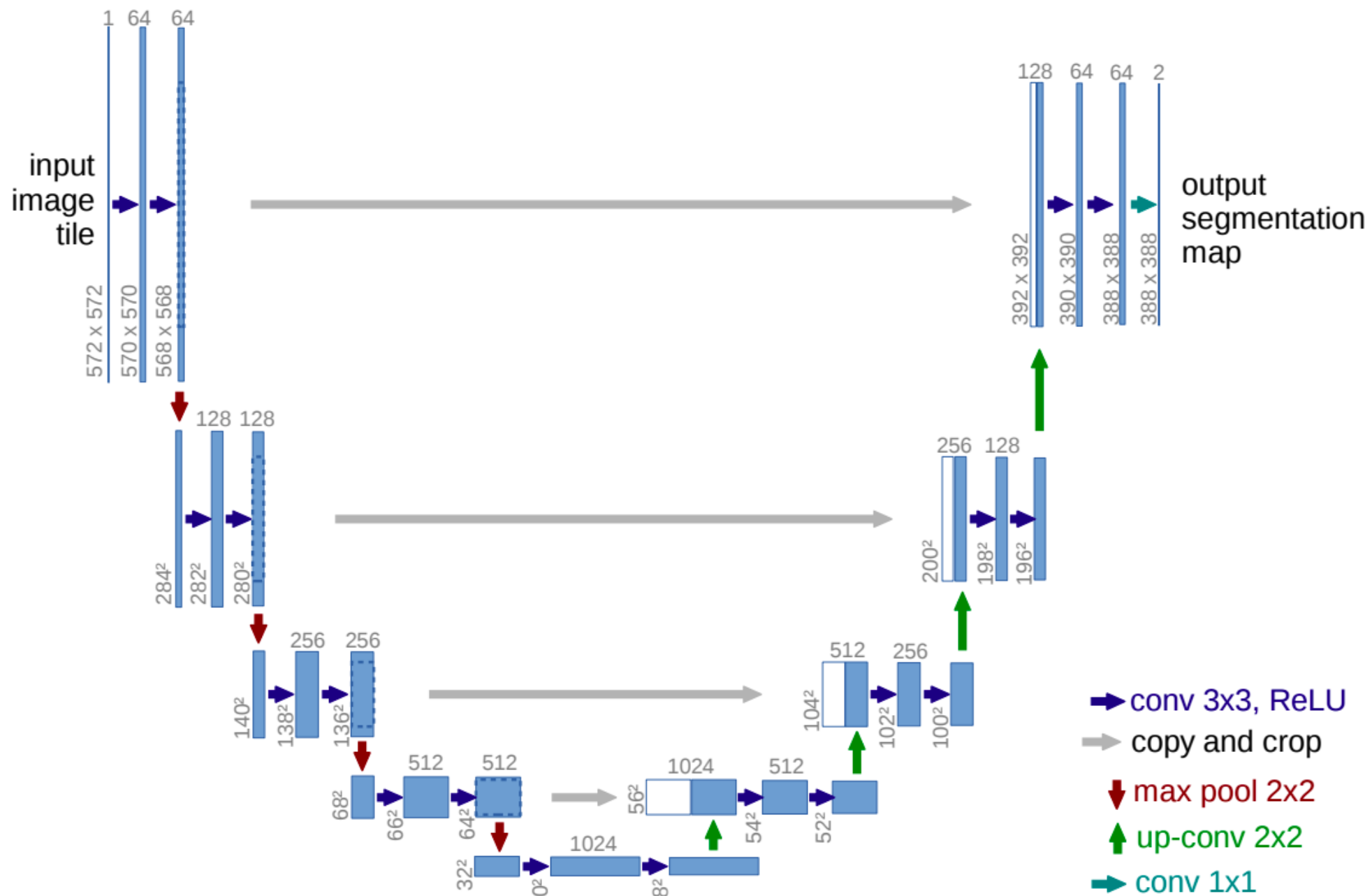
$$x_{23} = w_{23}z_{11} + w_{22}z_{12} + w_{13}z_{21} + w_{12}z_{22}$$

# Transpose Convolution



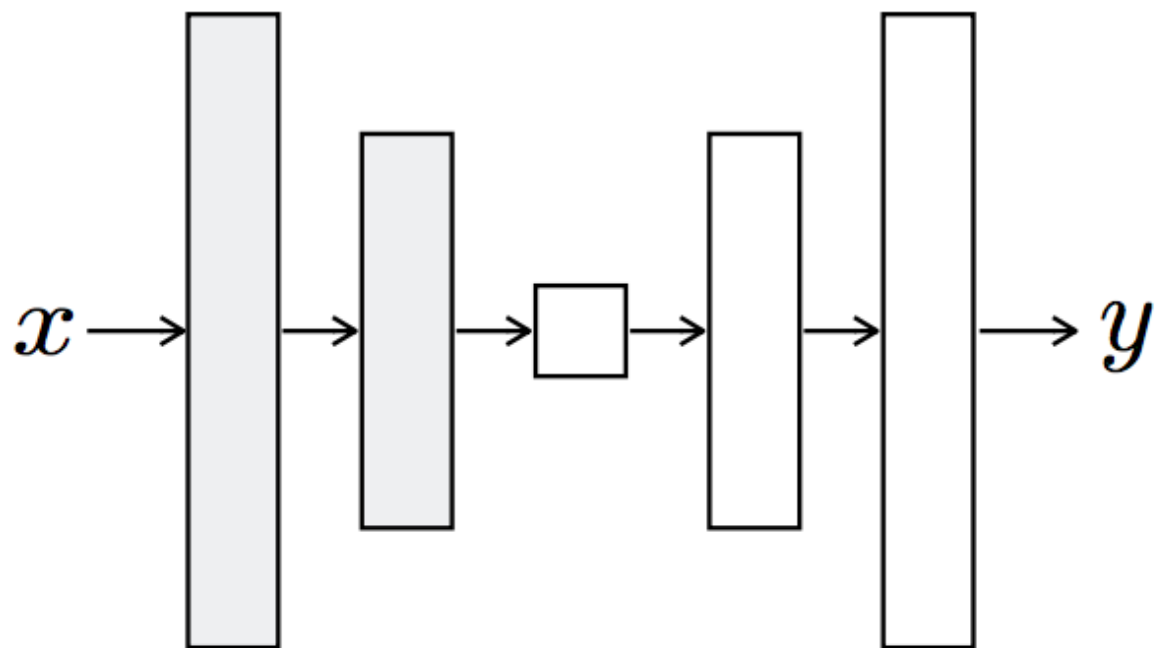
# Advanced Techniques in Segmentation

# U-Net

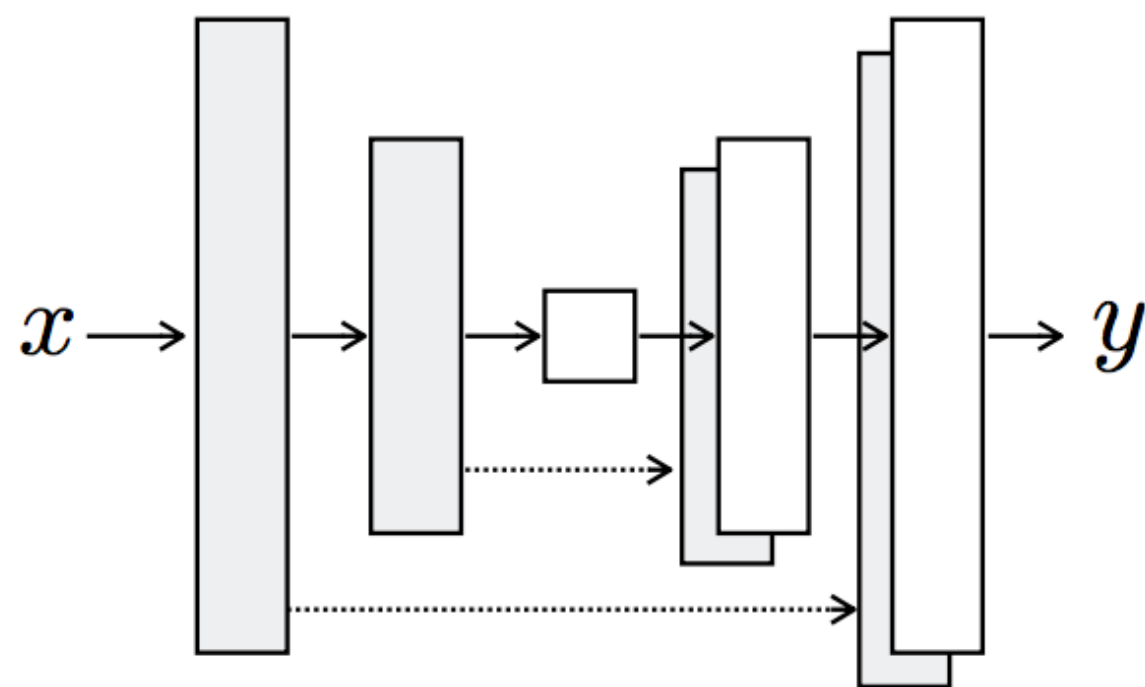


# U-Net

Encoder-decoder



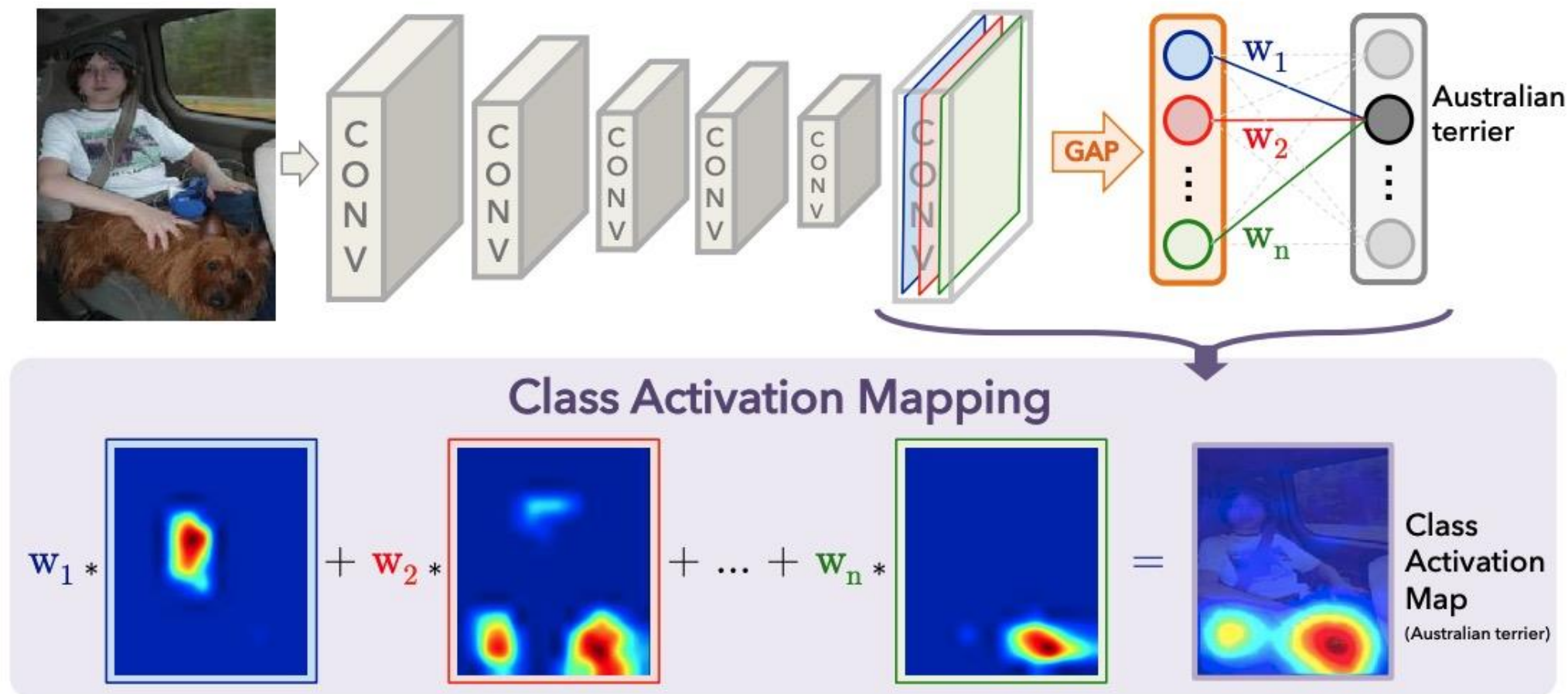
U-Net



# Visualizing Deep Networks using “Saliency map”



# CAM: Class Activation Maps



# CAM: Class Activation Maps

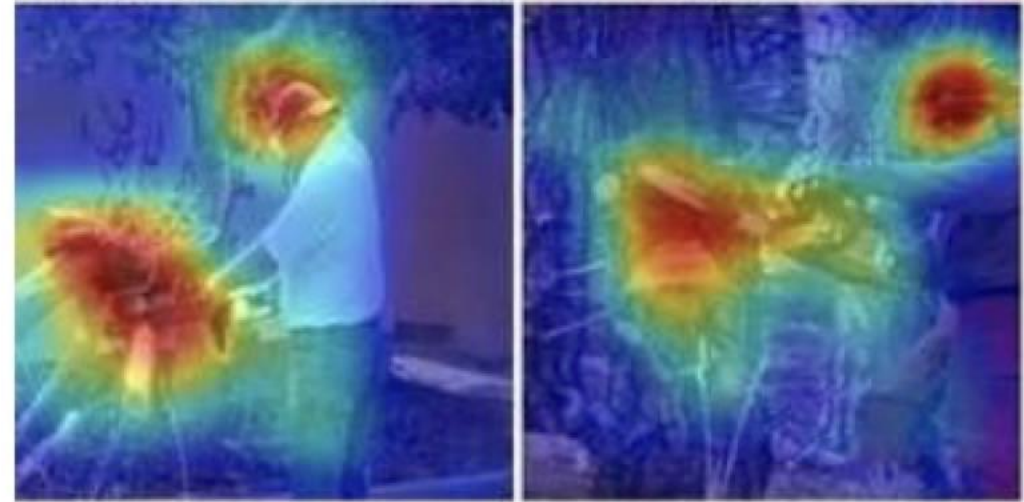


# CAM: Class Activation Maps

Brushing teeth

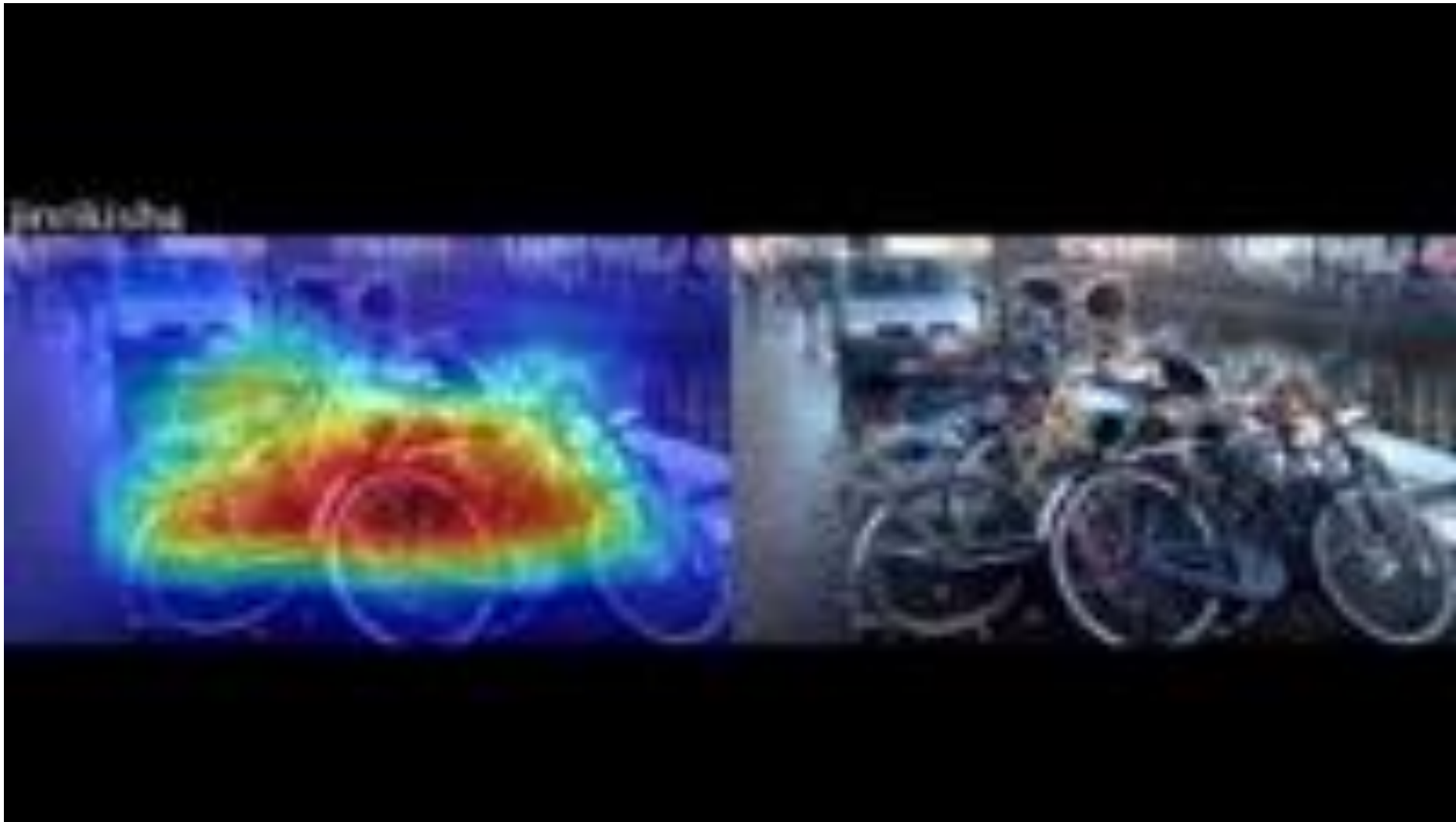


Cutting trees





# CAM: Class Activation Maps



# Visualizing Deep Networks by maximizing activation

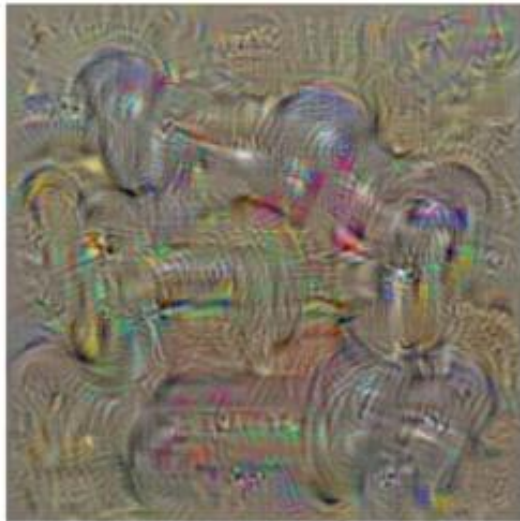
# Visualization by optimization

- We can synthesize images that maximize activation of a given neuron.
- Find image  $x$  maximizing target activation  $f(x)$  subject to *natural image regularization penalty*  $R(x)$ :

$$x^* = \arg \max_x f(x) - \lambda R(x)$$

# Visualization by optimization

- Maximize  $f(x) - \lambda R(x)$ 
  - $f(x)$  is score for a category *before softmax*
  - $R(x)$  is L2 regularization
  - Perform *gradient ascent* starting with zero image, add dataset mean to result



**dumbbell**



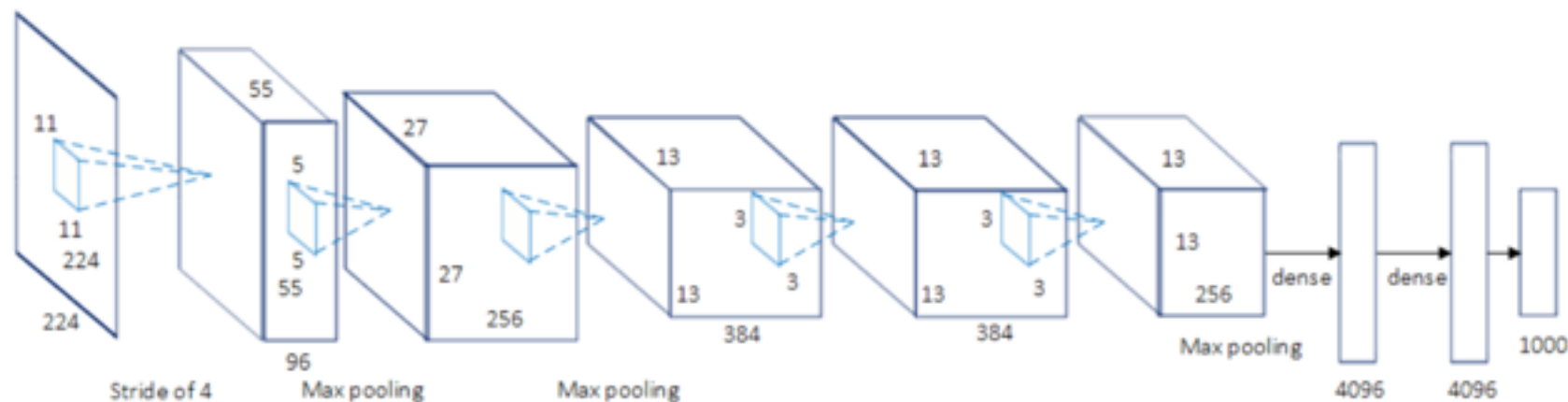
**cup**



**dalmatian**



# Visualization by optimization

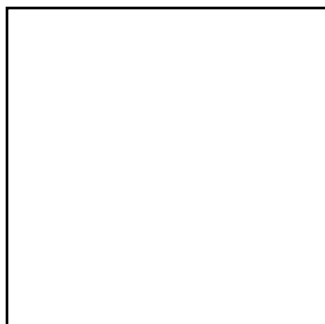


Compute the Loss



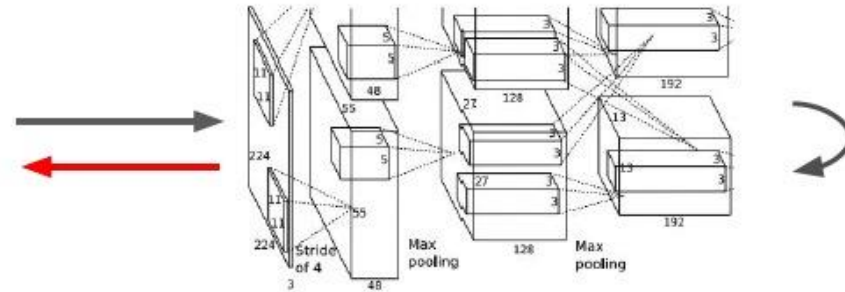
BP the gradients, but do not train the network

Keep adding/aggregating the gradients under constraint  $R(x)$



# Google DeepDream

Amplify one **layer** instead of just one neuron.



Choose an image and a layer in a CNN; repeat:

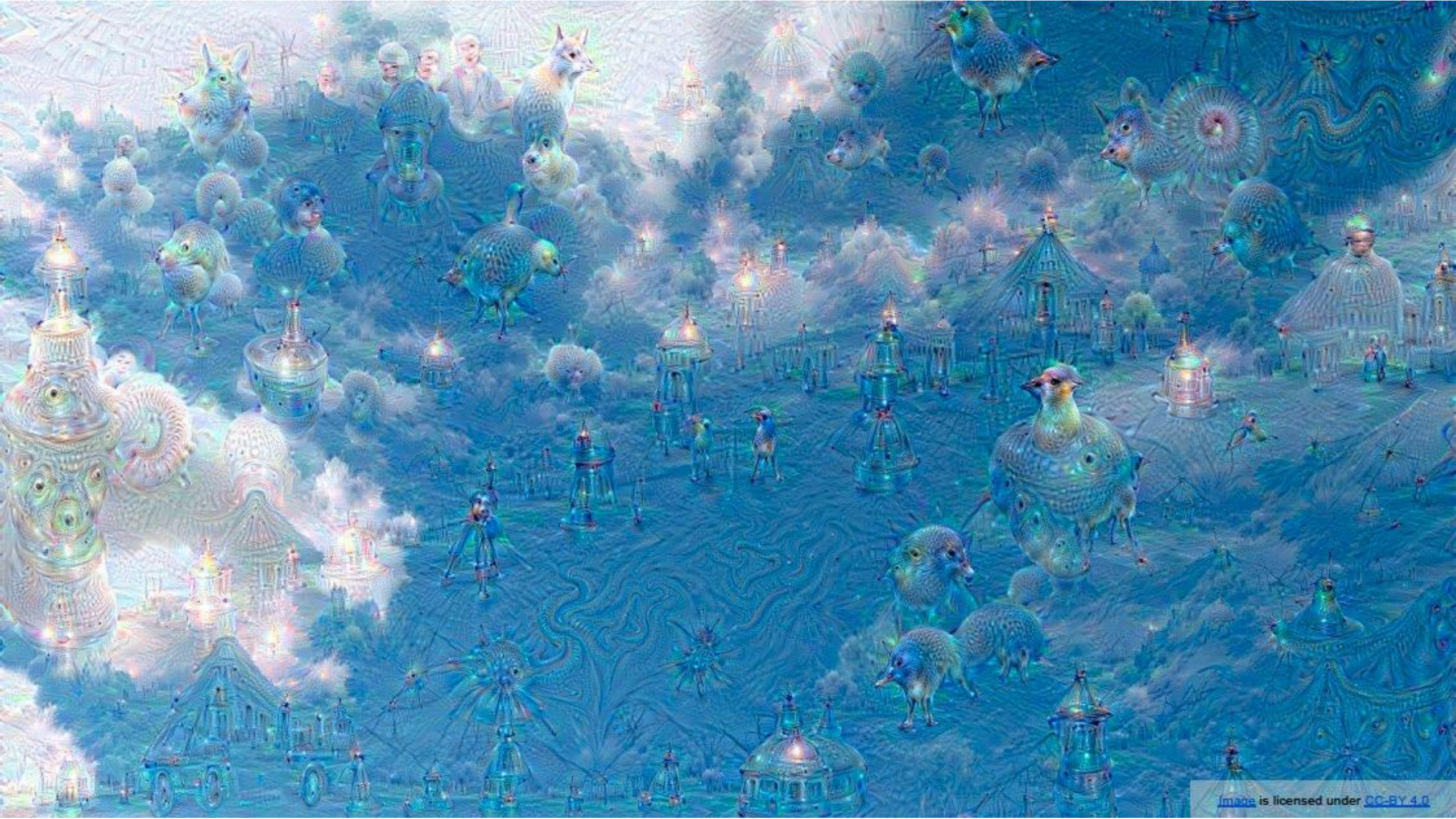
1. Forward: compute activations at chosen layer
2. Set gradient of chosen layer *equal to its activation*  
Equivalent to maximizing  $\sum_i f_i^2(x)$
3. Backward: Compute gradient w.r.t. image
4. Update image (with some tricks)



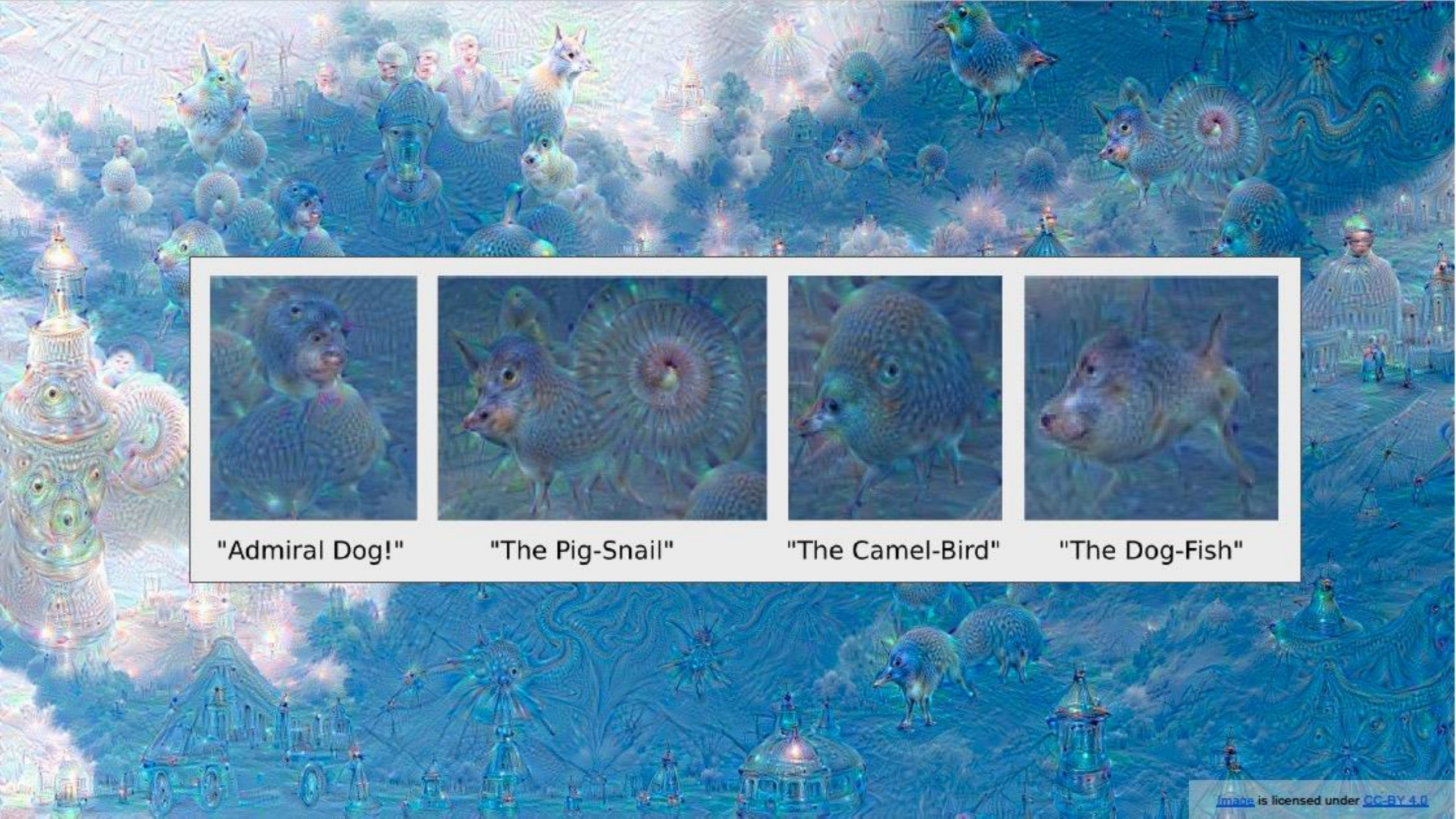












"Admiral Dog!"



"The Pig-Snail"



"The Camel-Bird"

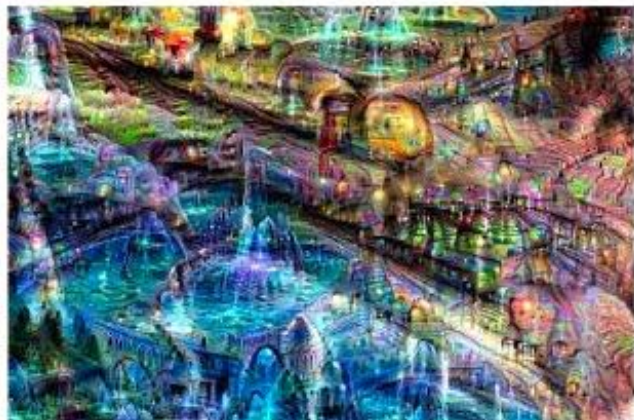
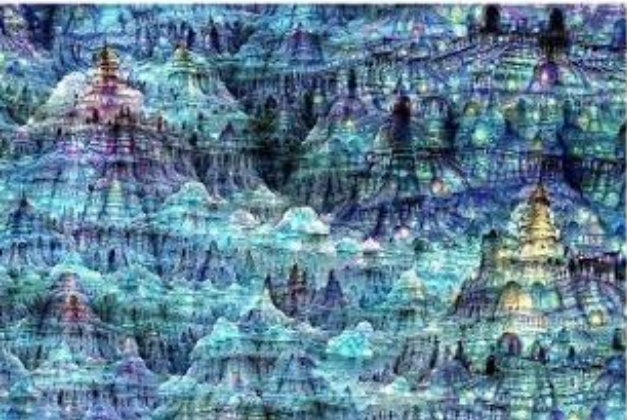


"The Dog-Fish"











# Next Class

Object Detection