Learning for vision III
Convolutional networks

Karel Zimmermann
http://cmp.felk.cvut.cz/~zimmerk/

Vision for Robotics and Autonomous Systems
https://cyber.felk.cvut.cz/vras/

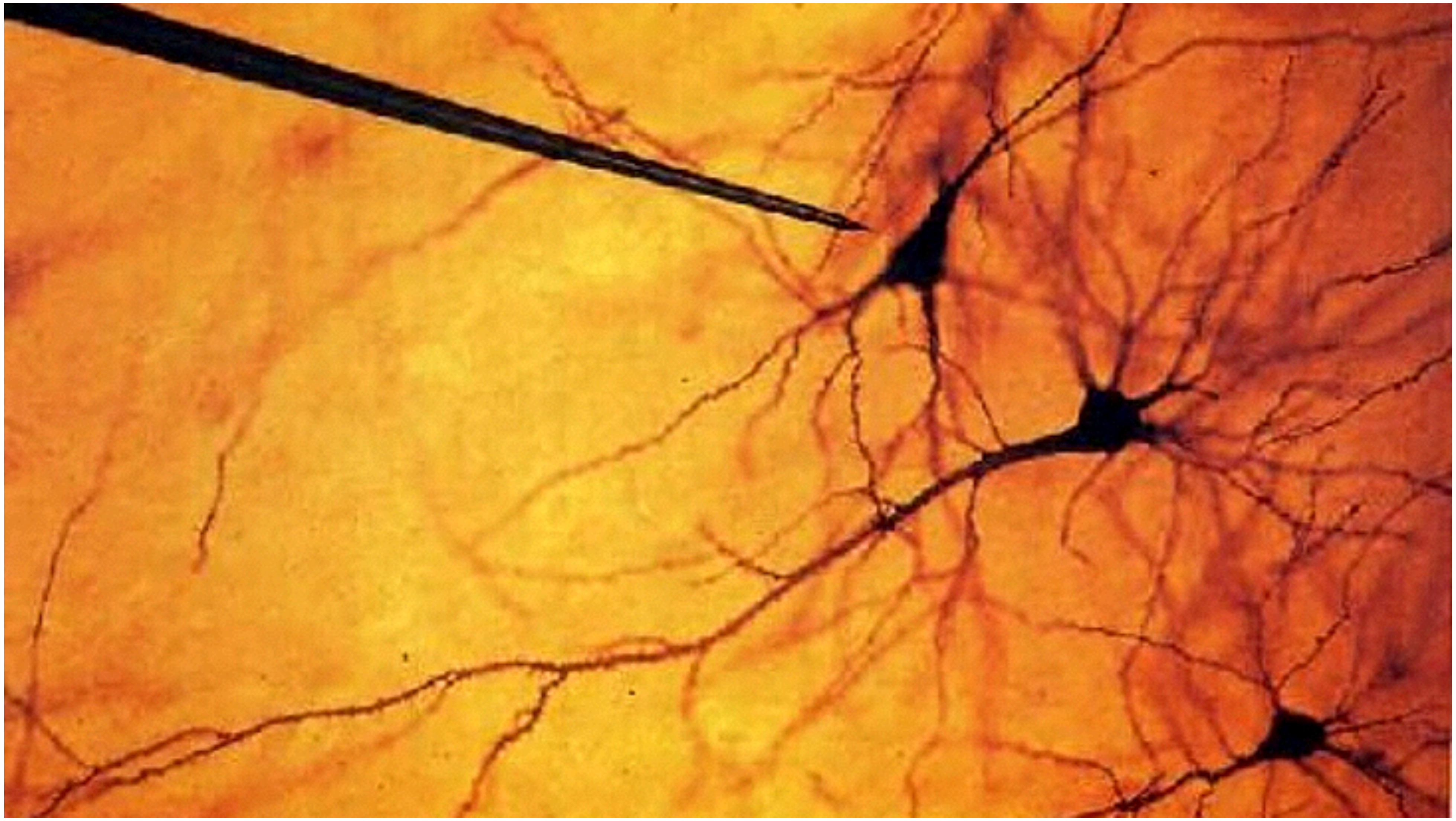
Center for Machine Perception
https://cmp.felk.cvut.cz

Department for Cybernetics
Faculty of Electrical Engineering
Czech Technical University in Prague

Outline

- Avoid overfitting by search for the NN model suitable for image processing [Hubel and Wiesel 1960].
- Feedforward and Backprop in ConvNets.

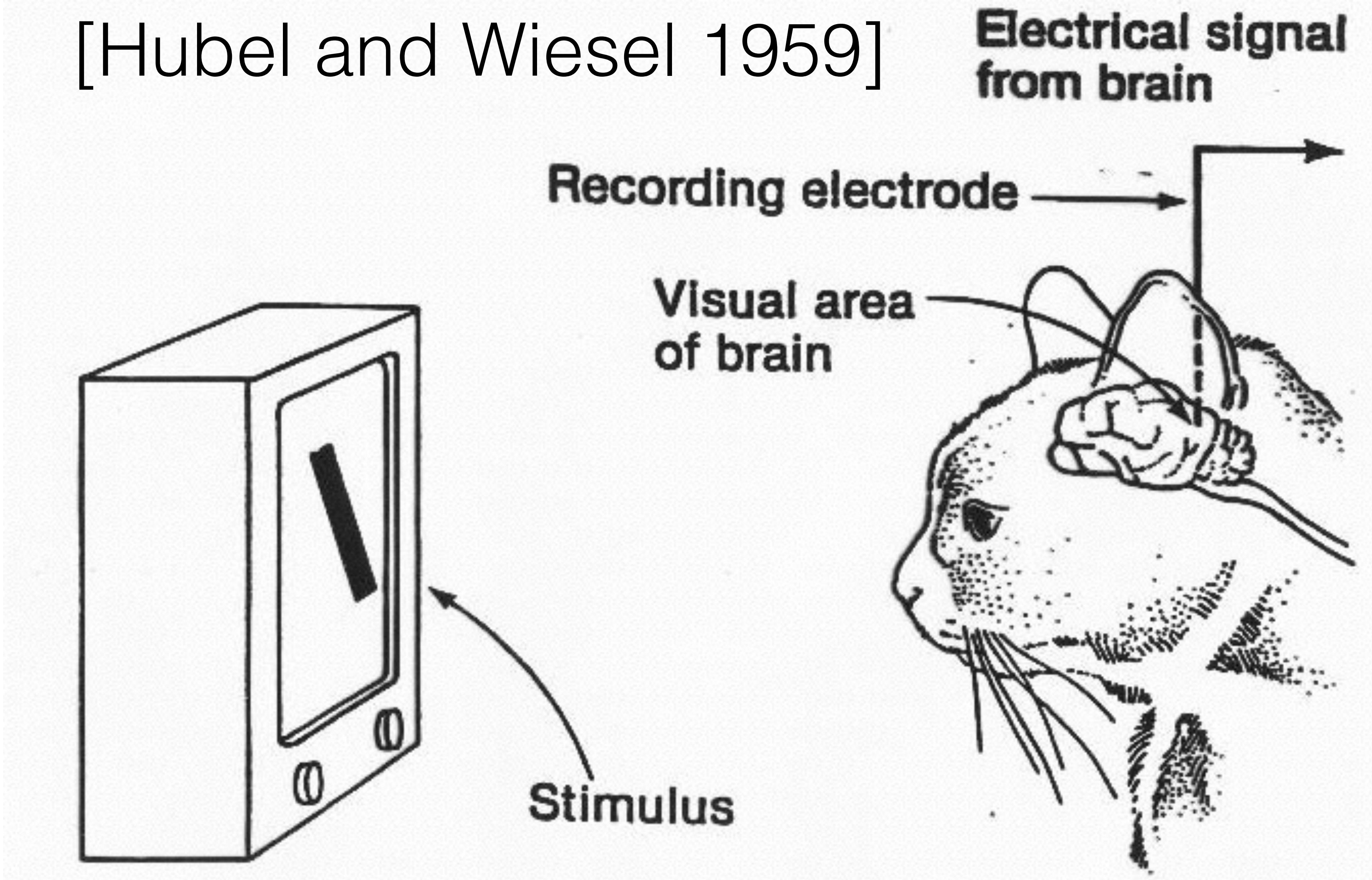
The Tungsten Electrode [Hubel-Science-1957]



http://braintour.harvard.edu/archives/portfolio-items/hubel-and-wiesel

- Device capable to record signal from a single neuron

[Hubel and Wiesel 1959]

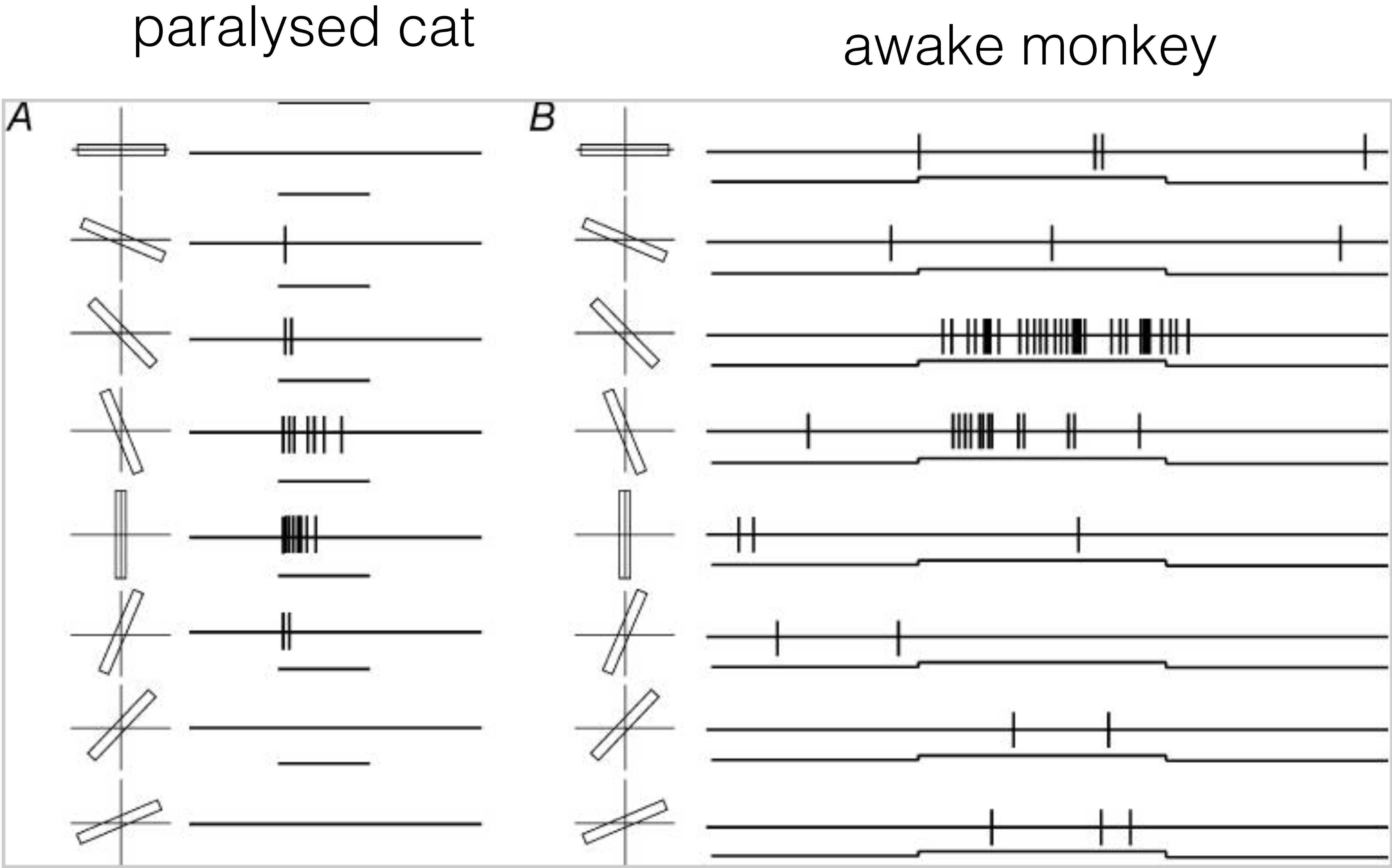


- Experiment with anaesthetised paralysed cat
- Recording of electrical signal reveals:
    1. Nearby neurons process information from nearby visual fields (topographical map).
    2. Neurons with similar function organized into columns
    3. Neurons are sensitive to edges and its orientation

[Hubel and Wiesel 1960]

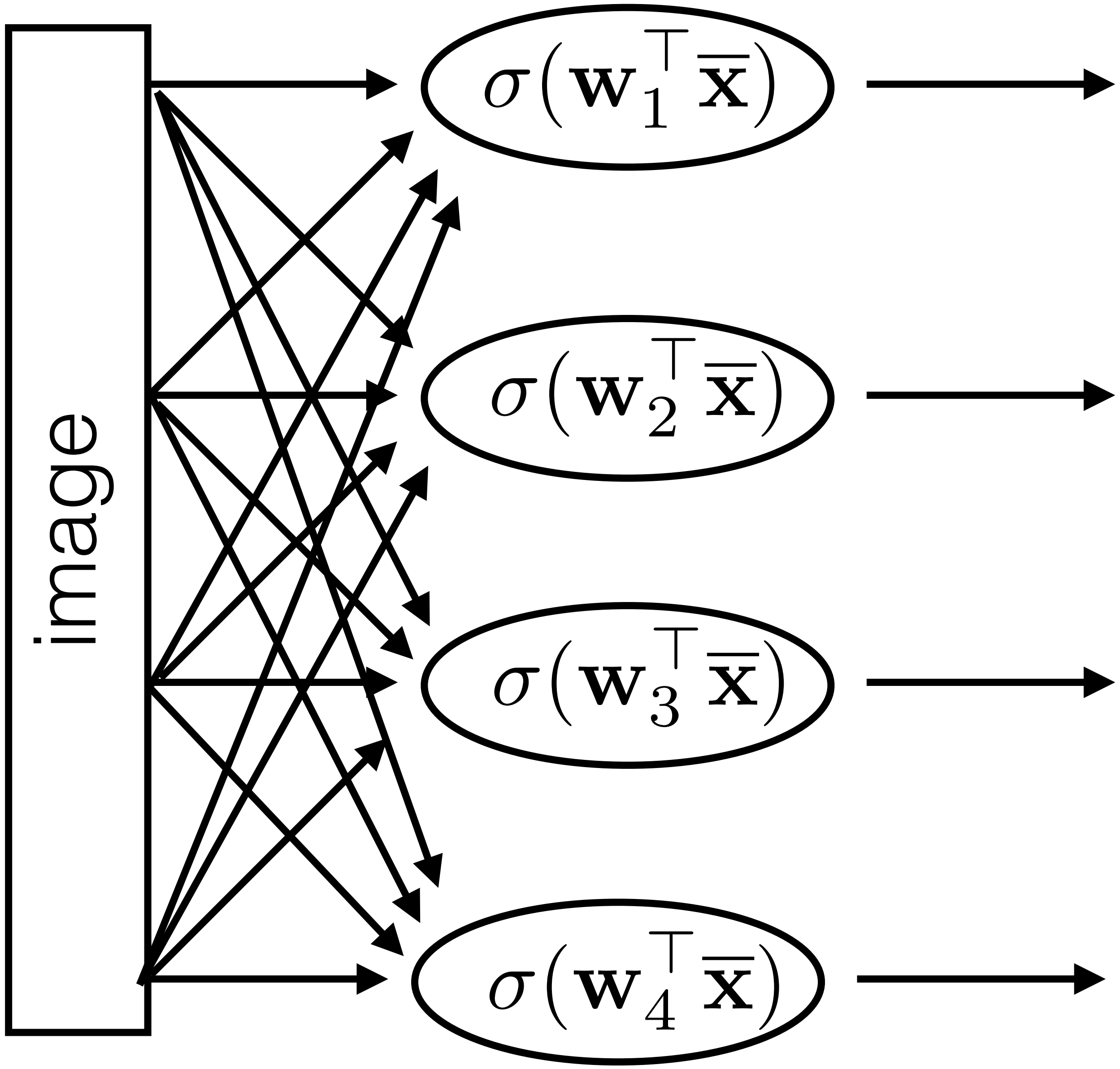
https://knowingneurons.com/2014/10/29/hubel-and-wiesel-the-neural-basis-of-visual-perception/

[Hubel and Wiesel 1960]



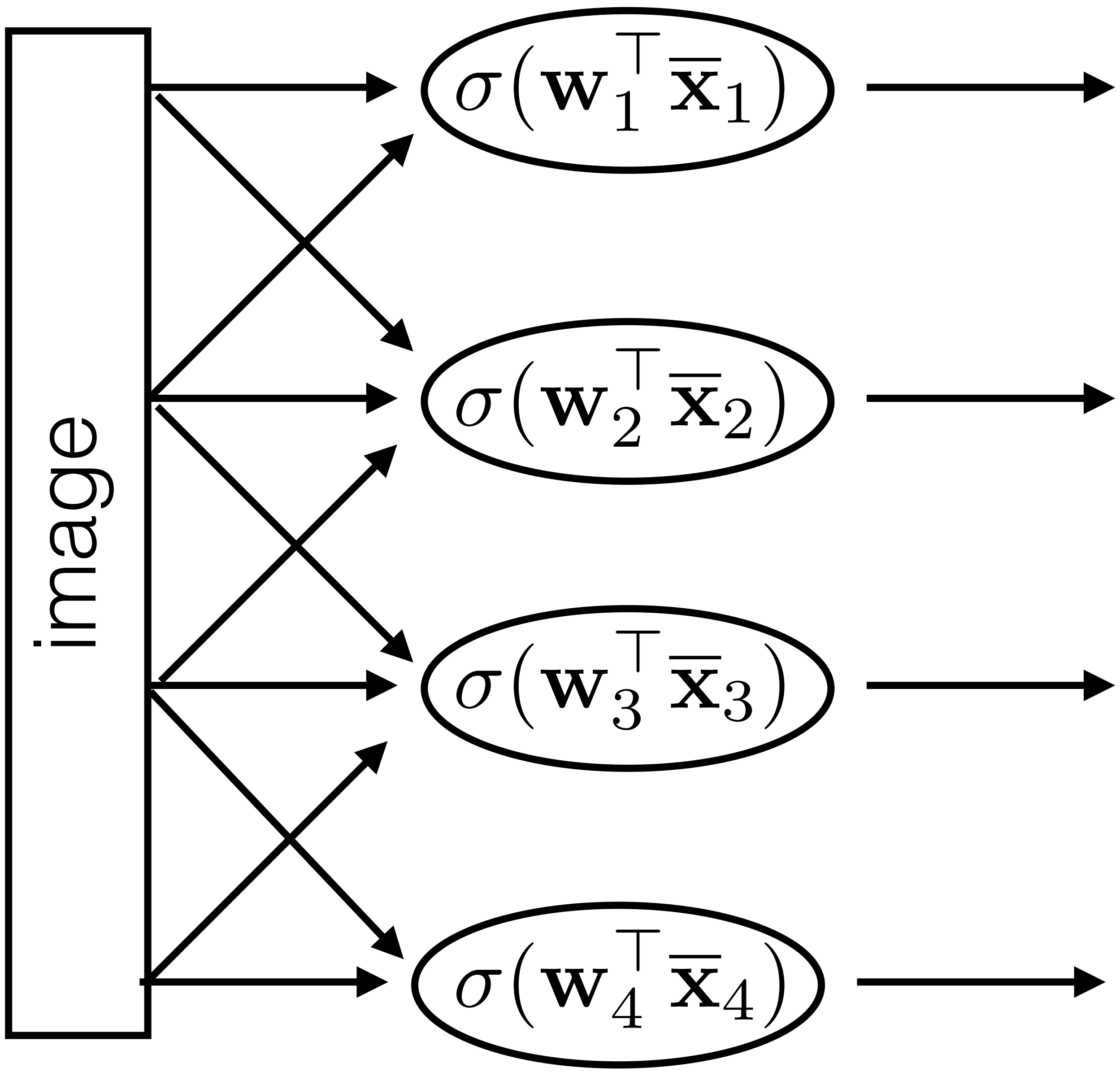
https://knowingneurons.com/2014/10/29/hubel-and-wiesel-the-neural-basis-of-visual-perception/

1. Nearby neurons process information from nearby visual field (topographical map).



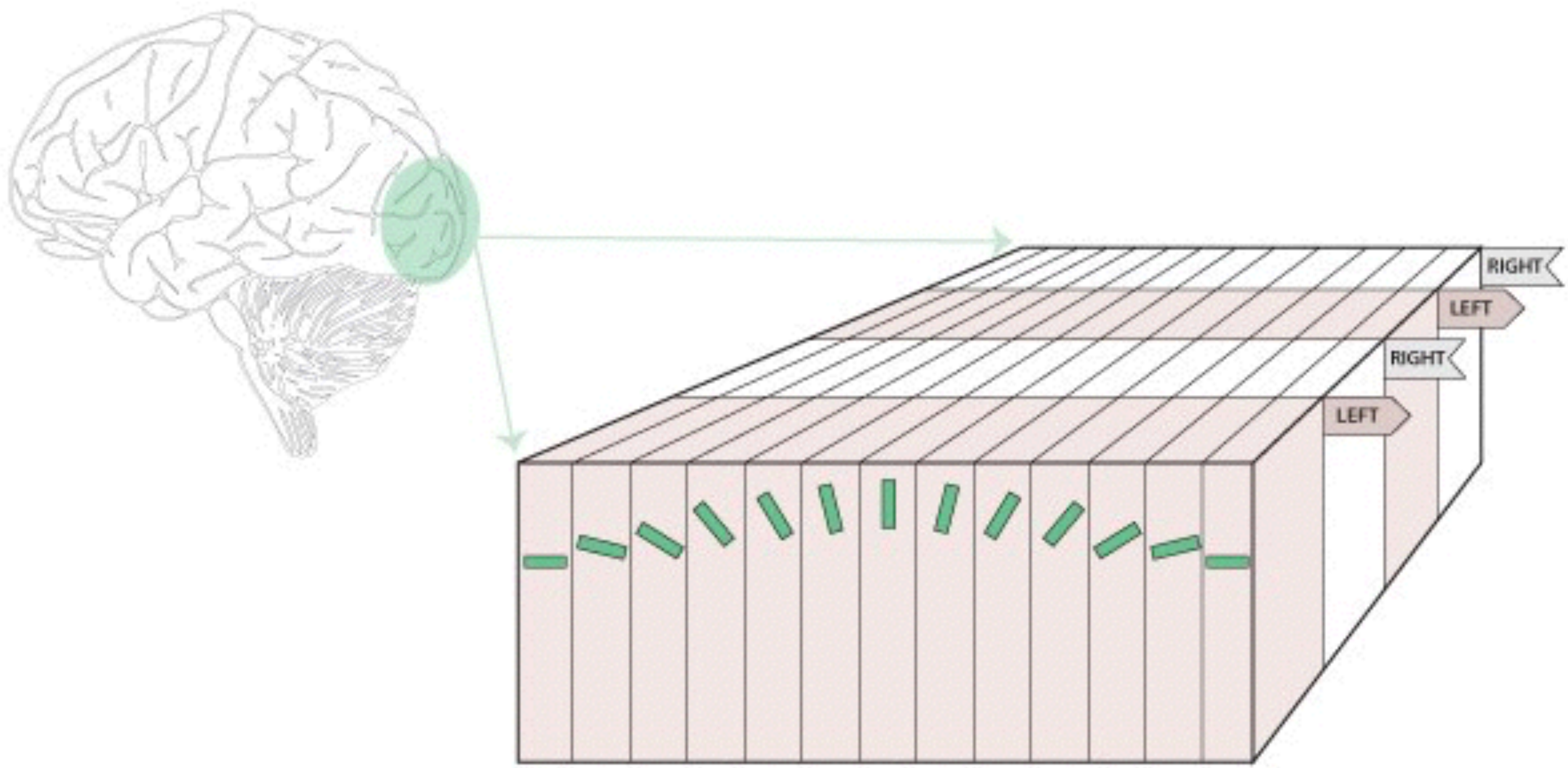
- Processing of visual information in cortex is not fully connected.

1. Nearby neurons process information from nearby visual field (topographical map).



- What is dimensionality reduction for N-pixel image and n-dimensional spatial neighbourhood?

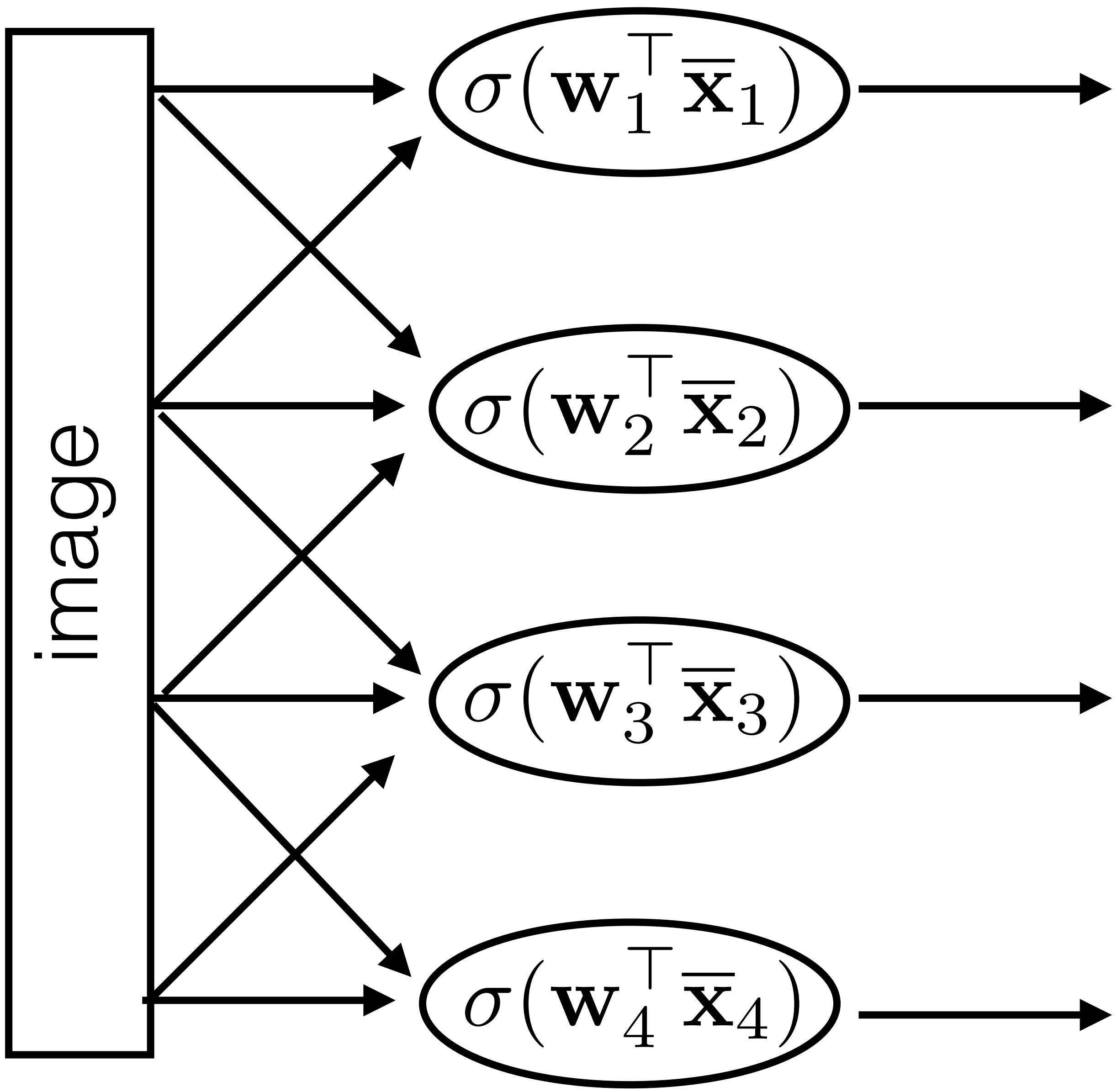
2. Neurons with similar function organized into columns



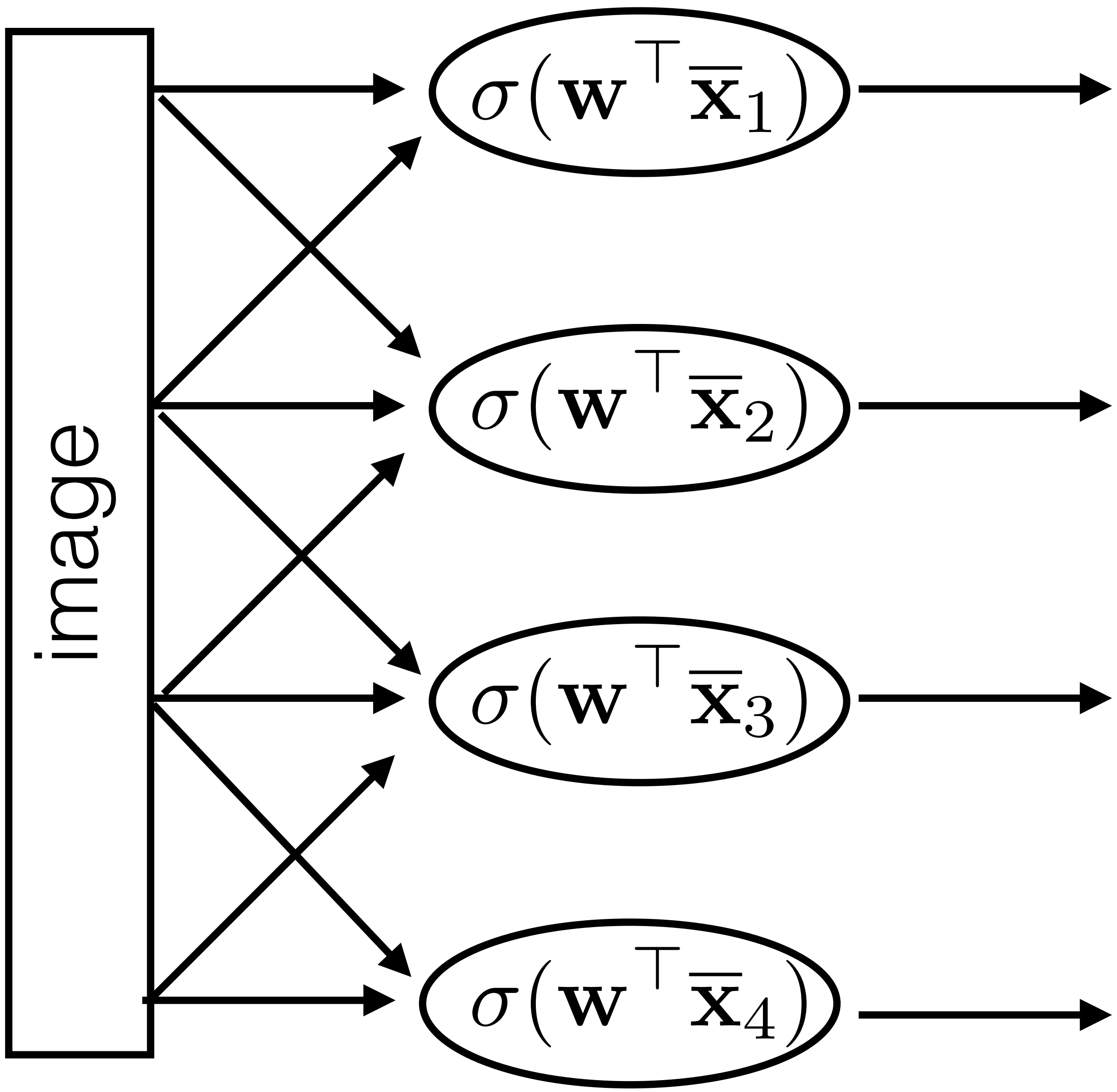
- There are neurons which detect an edge on the left and there are different neurons which detect the same edge on the right

https://knowingneurons.com/2014/10/29/hubel-and-wiesel-the-neural-basis-of-visual-perception/

2. Neurons with similar function organized into columns

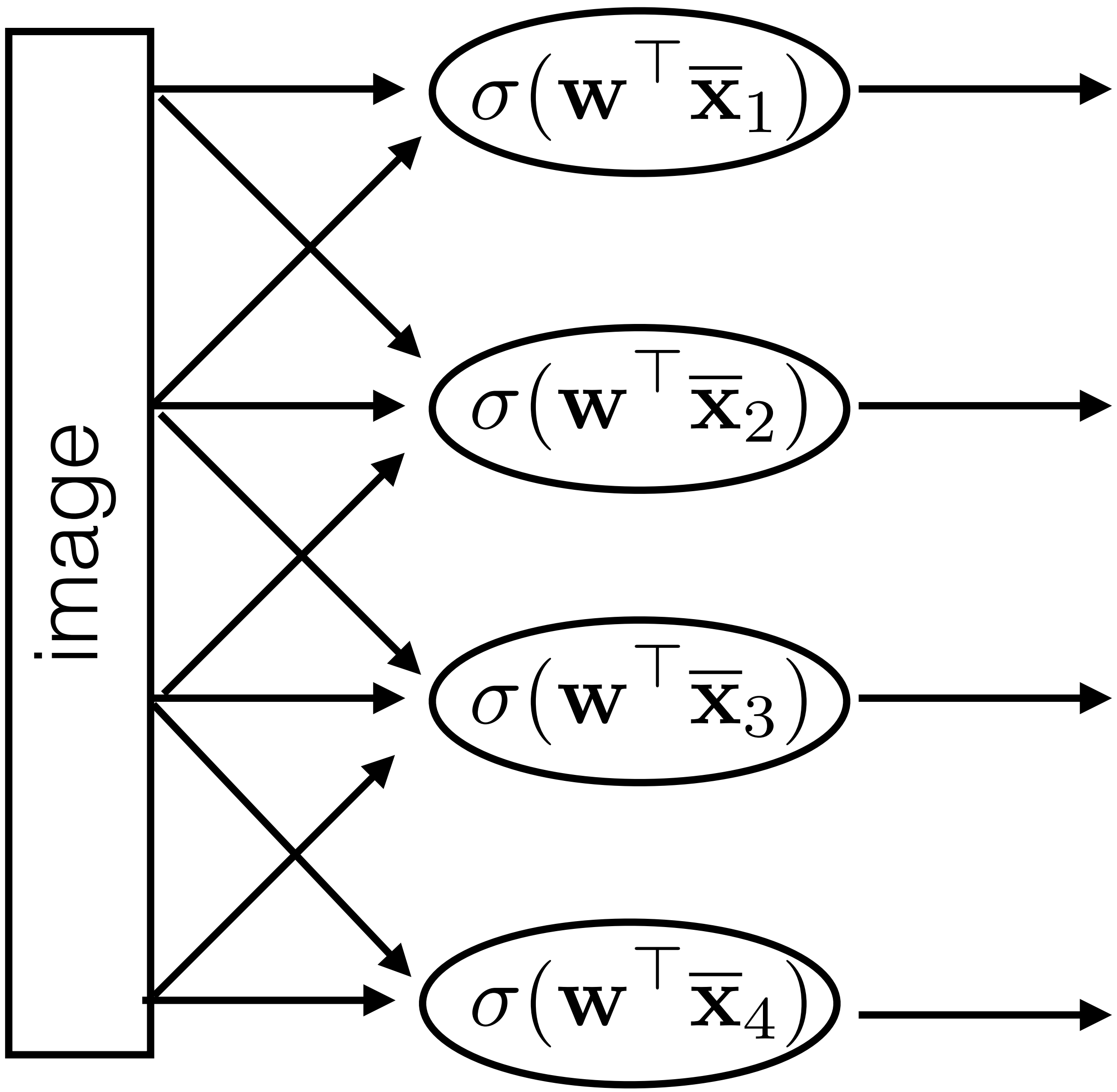


2. Neurons with similar function organized into columns



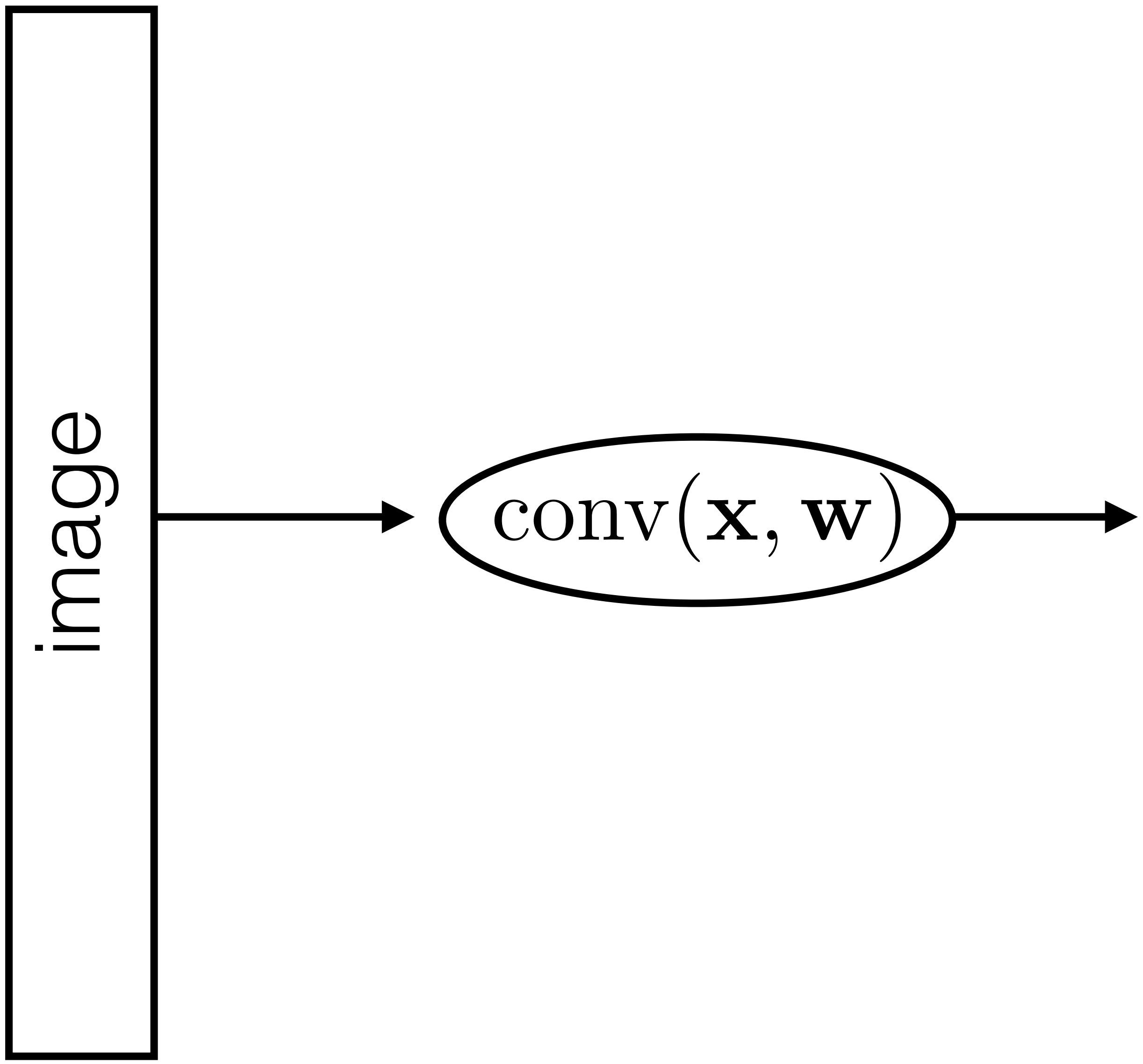
- What is dimensionality reduction for N pixel image and n-dimensional spatial neighbourhood?

2. Neurons with similar function organized into columns



- N^2 vs n (it does not depend on the image resolution)
- It corresponds to convolution of image $\mathbf{x}$ with kernel $\mathbf{w}$ followed by activation function

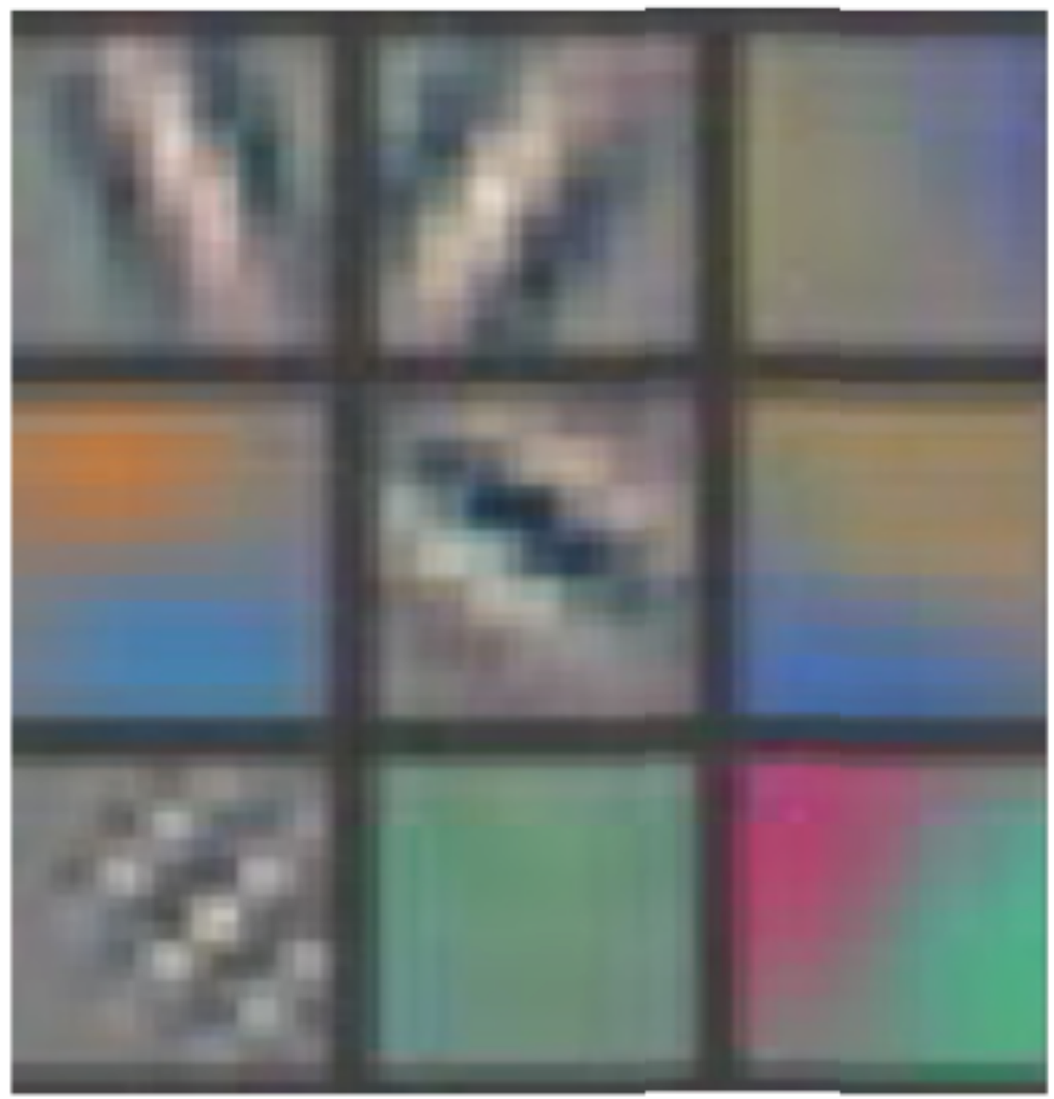
2. Neurons with similar function organized into columns



- N^2 vs n (it does not depend on the image resolution)
- It corresponds to convolution of image $\mathbf{x}$ with kernel $\mathbf{w}$ followed by activation function

3. Neurons are sensitive to edges and its orientation

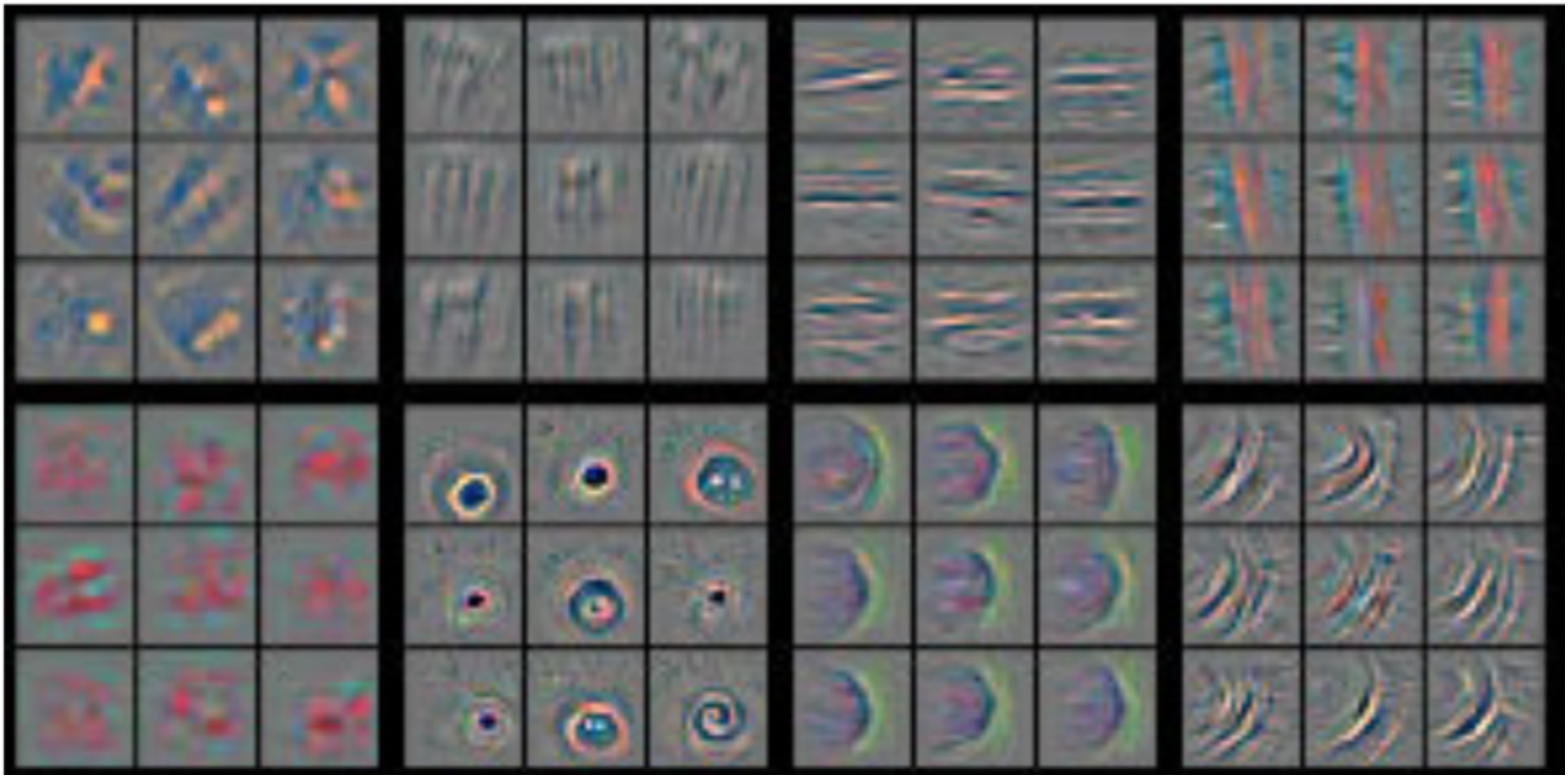
Inputs which maximized output of **layer 1**



[Zeiler and Fergus, ECCV, 2014]

3. Neurons are sensitive to edges and its orientation

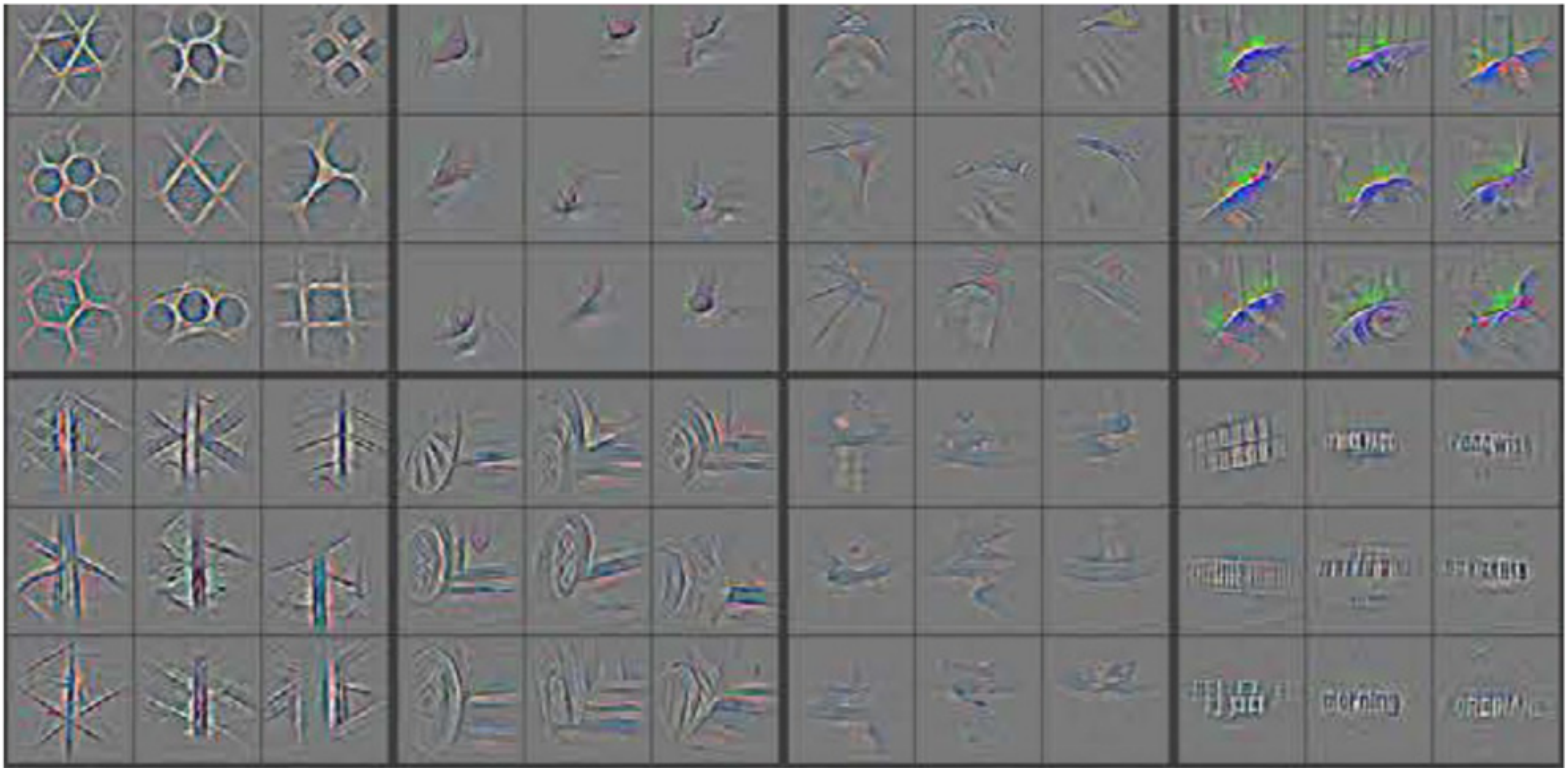
Inputs which maximized output of **layer 2**



[Zeiler and Fergus, ECCV, 2014]

3. Neurons are sensitive to edges and its orientation

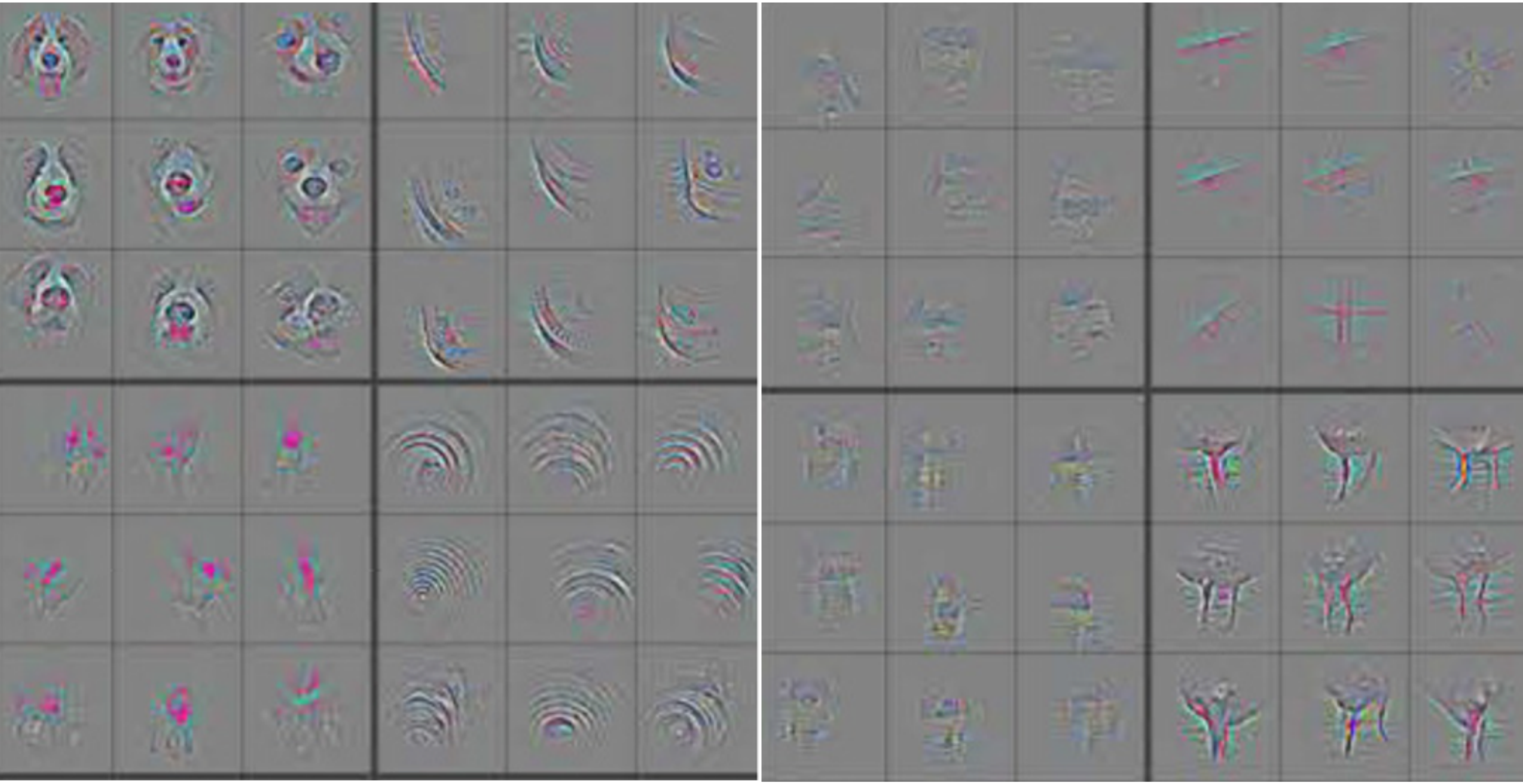
Inputs which maximized output of **layer 3**



[Zeiler and Fergus, ECCV, 2014]

3. Neurons are sensitive to edges and its orientation

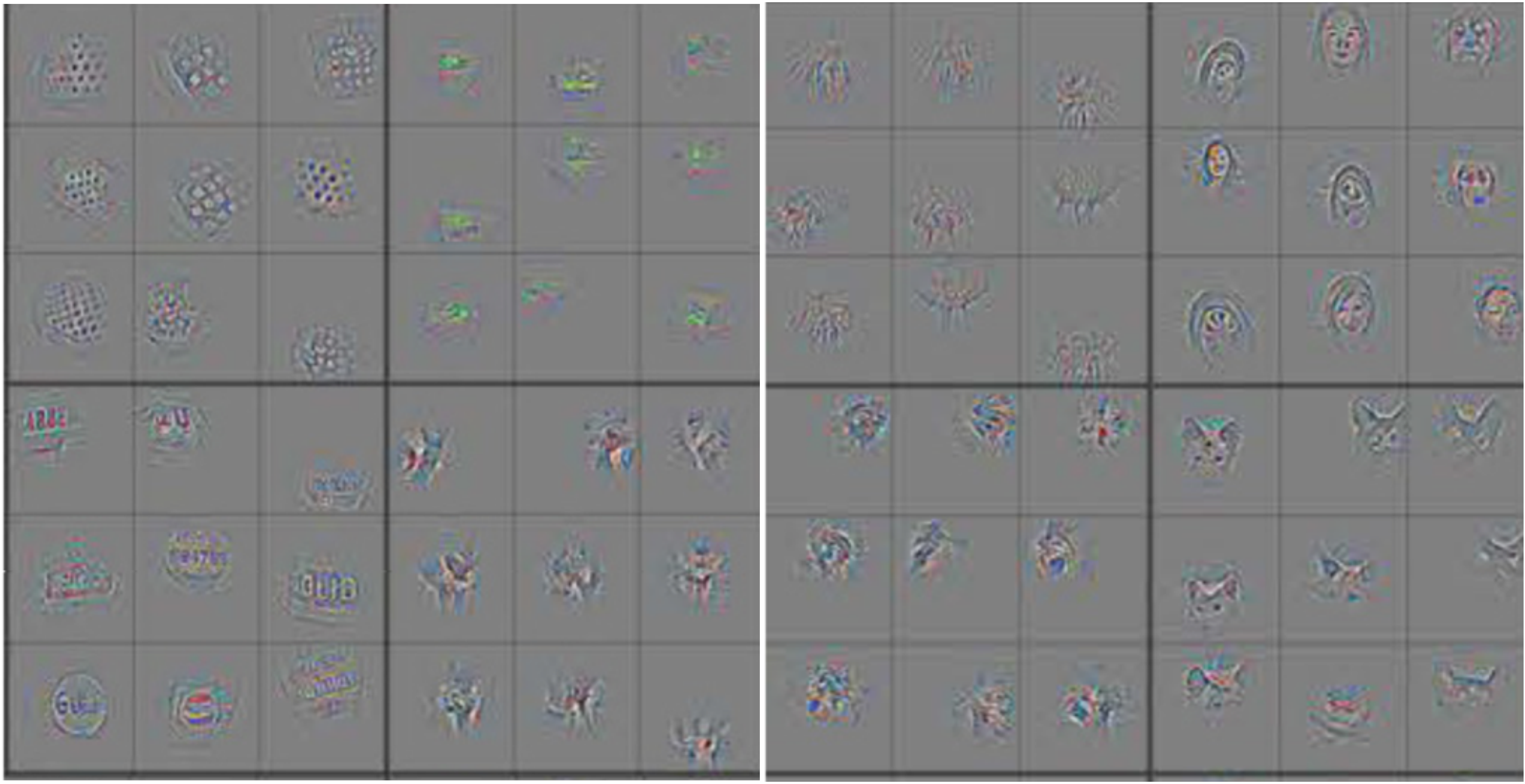
Inputs which maximized output of **layer 4**



[Zeiler and Fergus, ECCV, 2014]

3. Neurons are sensitive to edges and its orientation

Inputs which maximized output of **layer 5**



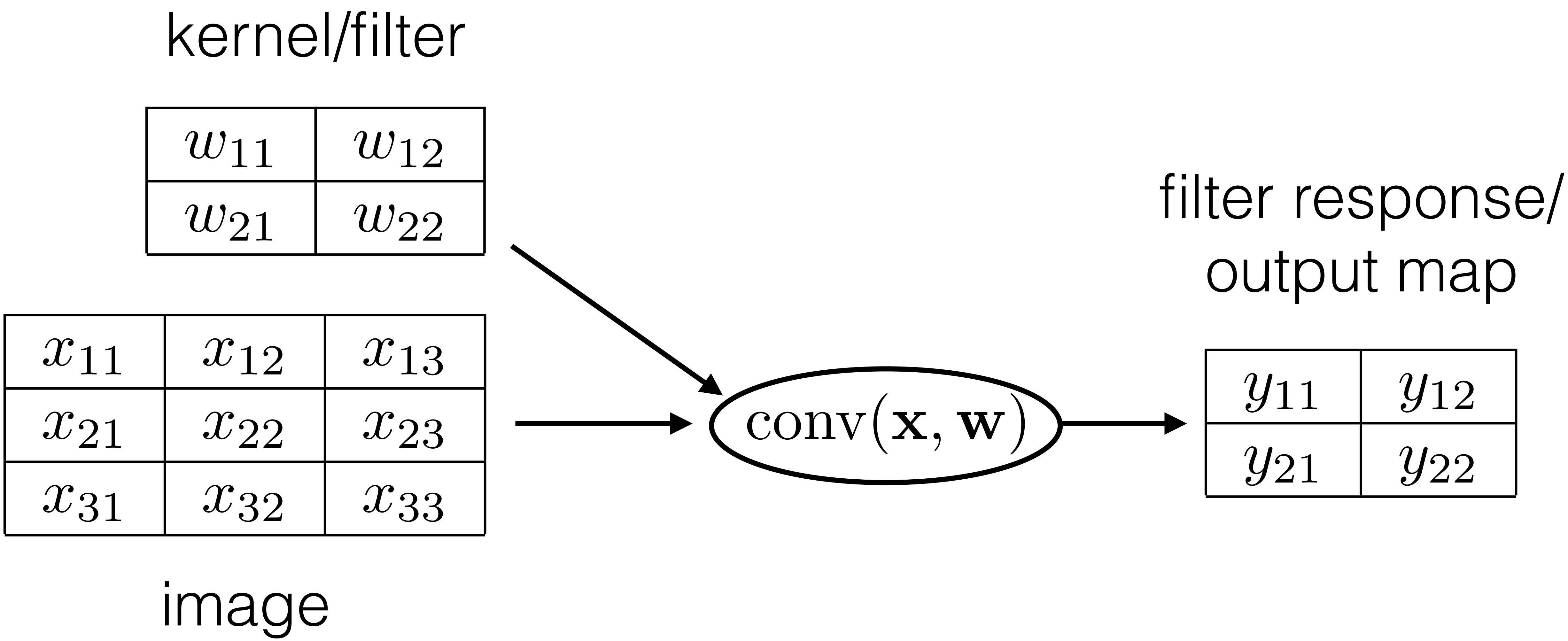
[Zeiler and Fergus, ECCV, 2014]

Hubel and Wiesel experiments in 1950s and 1960s

- Nobel Prize in Physiology and Medicine in 1981
- Dr. Hubel: “There has been a myth that the brain cannot understand itself. It is compared to a man trying to lift himself by his own bootstraps. We feel that is nonsense. The brain can be studied just as the kidney can.”

https://knowingneurons.com/2014/10/29/hubel-and-wiesel-the-neural-basis-of-visual-perception/

Convolution forward pass $\mathbf{y} = \mathrm{conv}(\mathbf{x}, \mathbf{w})$



Convolution forward pass $\mathbf{y} = \mathrm{conv}(\mathbf{x}, \mathbf{w})$

$$\begin{array}{|c|c|}\hline y_{11} & y_{12} \\ \hline y_{21} & y_{22} \\ \hline \end{array} = \mathrm{conv}\left(\begin{array}{|c|c|c|}\hline x_{11} & x_{12} & x_{13} \\ \hline x_{21} & x_{22} & x_{23} \\ \hline x_{31} & x_{32} & x_{33} \\ \hline \end{array}, \begin{array}{|c|c|}\hline w_{11} & w_{12} \\ \hline w_{21} & w_{22} \\ \hline \end{array} \right)$$

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

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

$$y_{21} = w_{11}x_{21} + w_{12}x_{22} + w_{21}x_{31} + w_{22}x_{32}$$

$$y_{22} = w_{11}x_{22} + w_{12}x_{23} + w_{21}x_{32} + w_{22}x_{33}$$

Convolution forward pass $\mathbf{y} = \mathrm{conv}(\mathbf{x}, \mathbf{w})$

$$\begin{array}{|c|c|}\hline y_{11} & y_{12} \\ \hline y_{21} & y_{22} \\ \hline \end{array} = \mathrm{conv}\left(\begin{array}{|c|c|c|}\hline x_{11} & x_{12} & x_{13} \\ \hline x_{21} & x_{22} & x_{23} \\ \hline x_{31} & x_{32} & x_{33} \\ \hline \end{array}, \begin{array}{|c|c|}\hline w_{11} & w_{12} \\ \hline w_{21} & w_{22} \\ \hline \end{array} \right)$$

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

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

$$y_{21} = w_{11}x_{21} + w_{12}x_{22} + w_{21}x_{31} + w_{22}x_{32}$$

$$y_{22} = w_{11}x_{22} + w_{12}x_{23} + w_{21}x_{32} + w_{22}x_{33}$$

Convolution forward pass $\mathbf{y} = \text{conv}(\mathbf{x}, \mathbf{w})$

$$\begin{array}{|c|c|}\hline y_{11} & \boxed{y_{12}} \\ \hline y_{21} & y_{22} \\ \hline\end{array} = \text{conv}\left(\begin{array}{|c|c|c|}\hline x_{11} & x_{12} & x_{13} \\ \hline x_{21} & x_{22} & x_{23} \\ \hline x_{31} & x_{32} & x_{33} \\ \hline\end{array}, \begin{array}{|c|c|}\hline w_{11} & w_{12} \\ \hline w_{21} & w_{22} \\ \hline\end{array}\right)$$

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

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

$$y_{21} = w_{11}x_{21} + w_{12}x_{22} + w_{21}x_{31} + w_{22}x_{32}$$

$$y_{22} = w_{11}x_{22} + w_{12}x_{23} + w_{21}x_{32} + w_{22}x_{33}$$

Convolution forward pass $\mathbf{y} = \mathrm{conv}(\mathbf{x}, \mathbf{w})$

$$\begin{array}{|c|c|}\hline y_{11} & y_{12} \\ \hline y_{21} & y_{22} \\ \hline\end{array} = \mathrm{conv}\left(\begin{array}{|c|c|c|}\hline x_{11} & x_{12} & x_{13} \\ \hline x_{21} & x_{22} & x_{23} \\ \hline x_{31} & x_{32} & x_{33} \\ \hline\end{array}, \begin{array}{|c|c|}\hline w_{11} & w_{12} \\ \hline w_{21} & w_{22} \\ \hline\end{array} \right)$$

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

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

$$y_{21} = w_{11}x_{21} + w_{12}x_{22} + w_{21}x_{31} + w_{22}x_{32}$$

$$y_{22} = w_{11}x_{22} + w_{12}x_{23} + w_{21}x_{32} + w_{22}x_{33}$$

Convolution forward pass $\mathbf{y} = \mathrm{conv}(\mathbf{x}, \mathbf{w})$

$$\begin{array}{|c|c|}\hline y_{11} & y_{12} \\ \hline y_{21} & \boxed{y_{22}} \\ \hline \end{array} = \mathrm{conv}\left(\begin{array}{|c|c|c|}\hline x_{11} & x_{12} & x_{13} \\ \hline x_{21} & x_{22} & x_{23} \\ \hline x_{31} & x_{32} & x_{33} \\ \hline \end{array}, \begin{array}{|c|c|}\hline w_{11} & w_{12} \\ \hline w_{21} & w_{22} \\ \hline \end{array} \right)$$

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

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

$$y_{21} = w_{11}x_{21} + w_{12}x_{22} + w_{21}x_{31} + w_{22}x_{32}$$

$$\boxed{y_{22} = w_{11}x_{22} + w_{12}x_{23} + w_{21}x_{32} + w_{22}x_{33}}$$

Convolution layer properties - output size

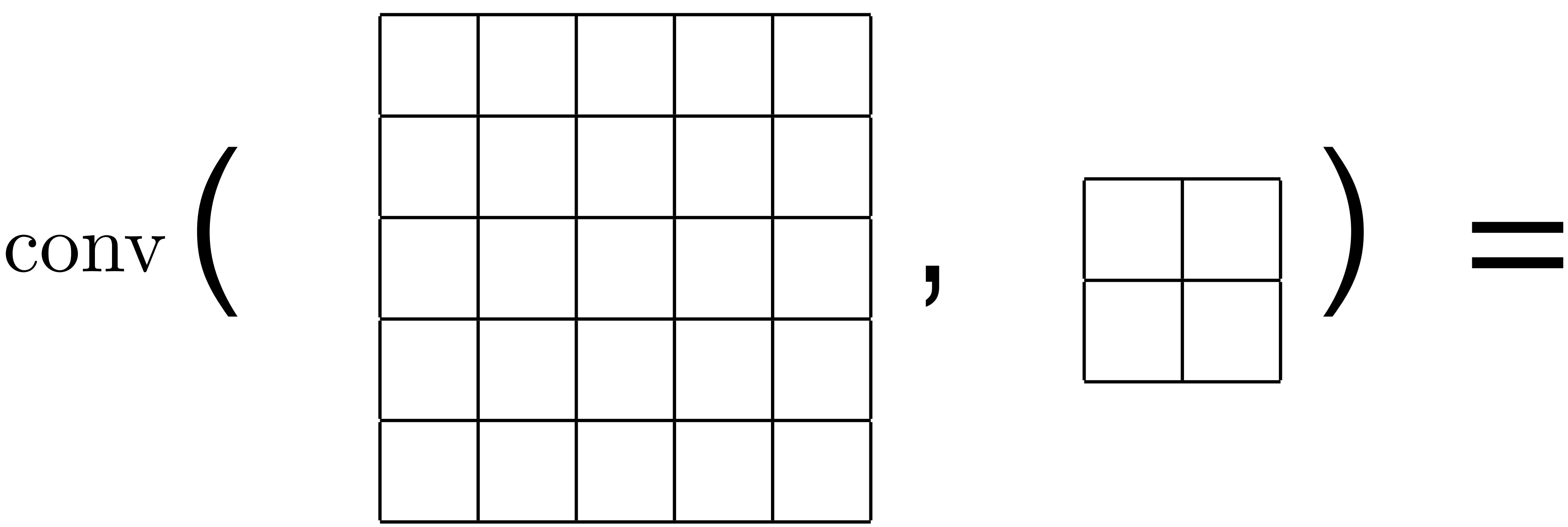
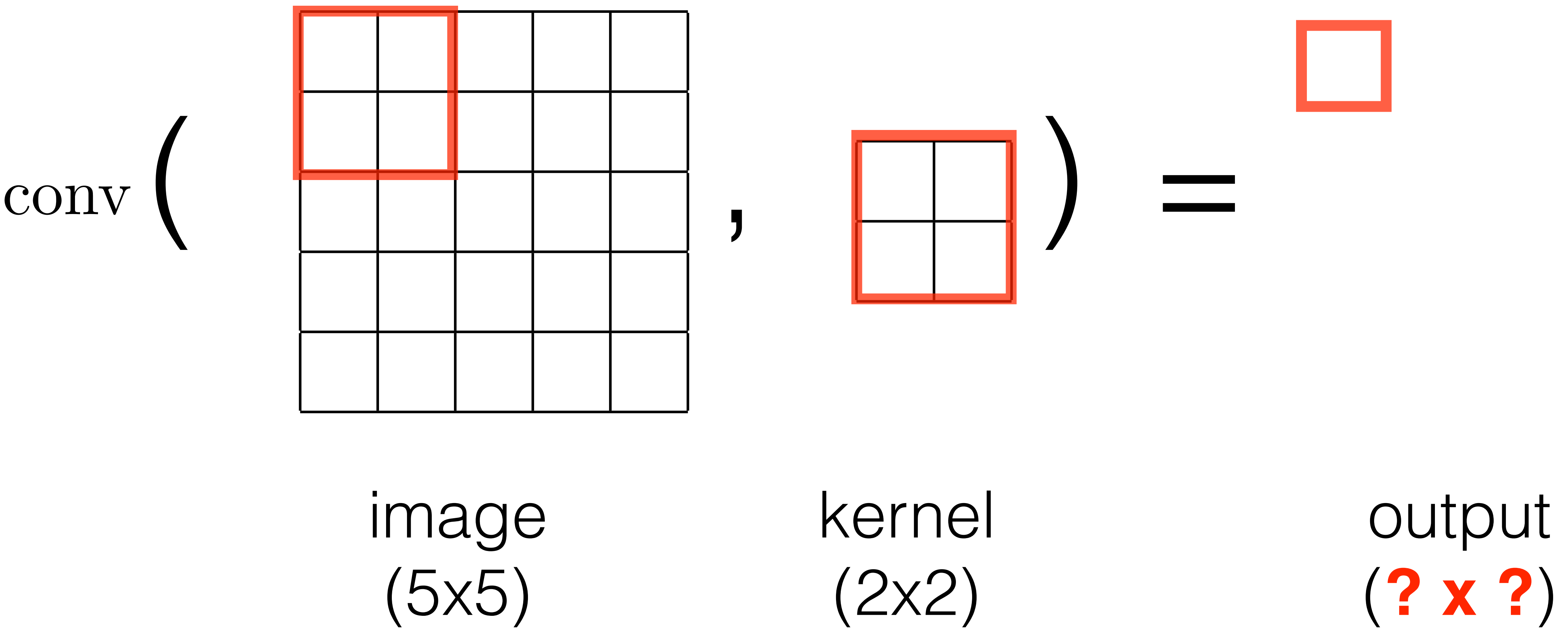


image
(5x5)

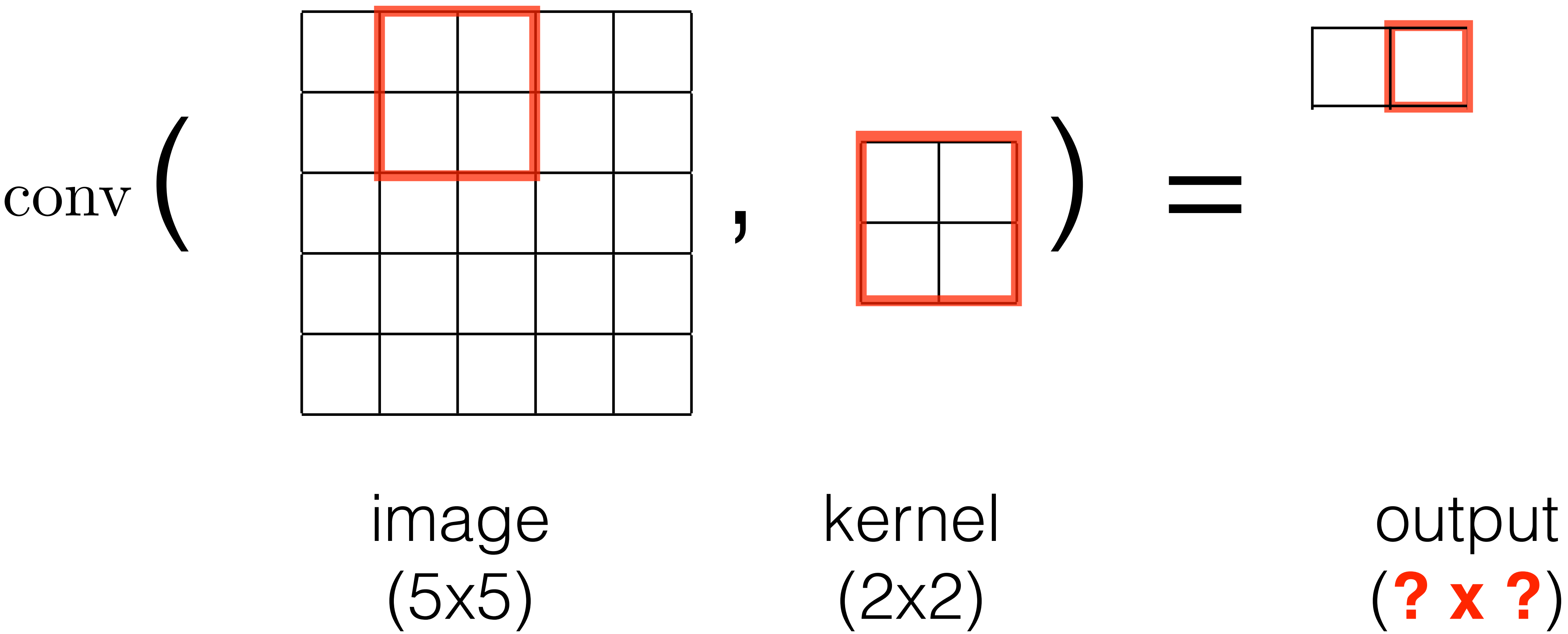
kernel
(2x2)

output
(**? x ?**)

Convolution layer properties - output size



Convolution layer properties - output size



Convolution layer properties - output size

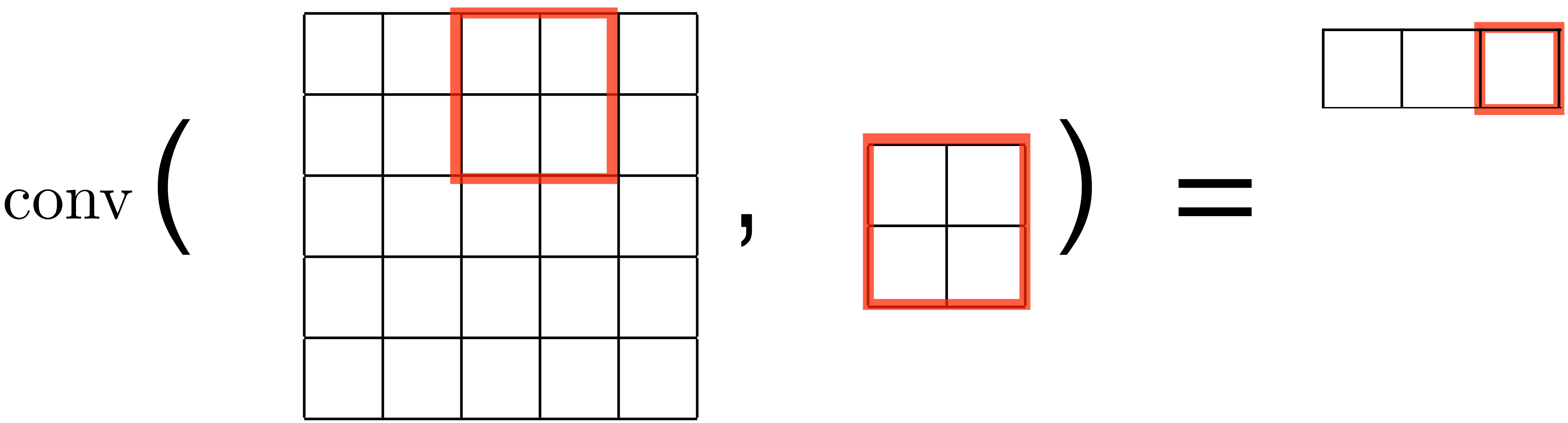


image
(5x5)

kernel
(2x2)

output
(**? x ?**)

Convolution layer properties - output size

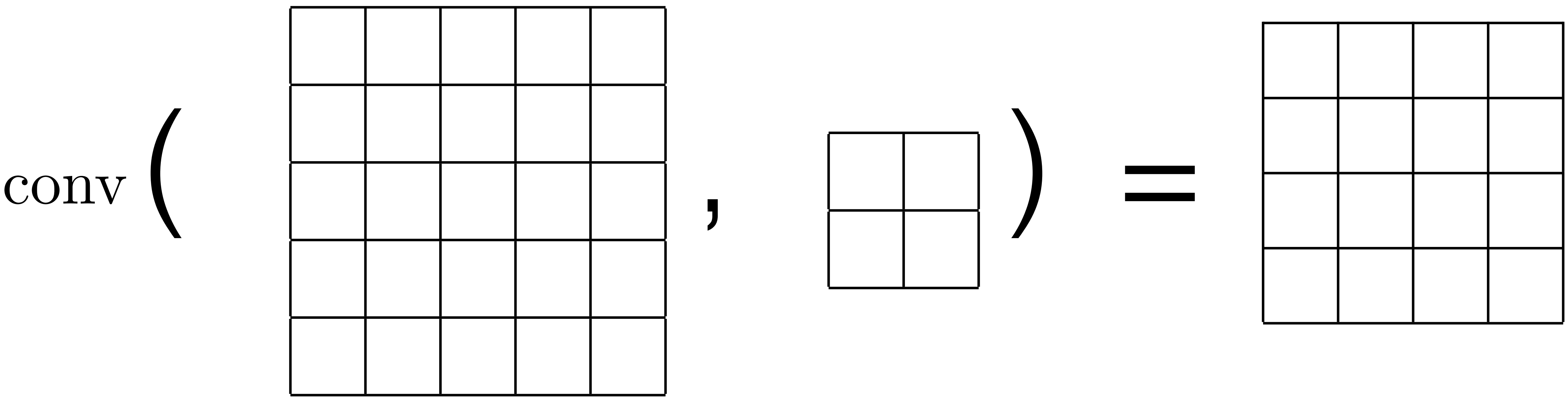


image (5x5) kernel (2x2) output (4x4)

Convolution layer properties - stride

stride = 1

kernel moves by 1 pixel

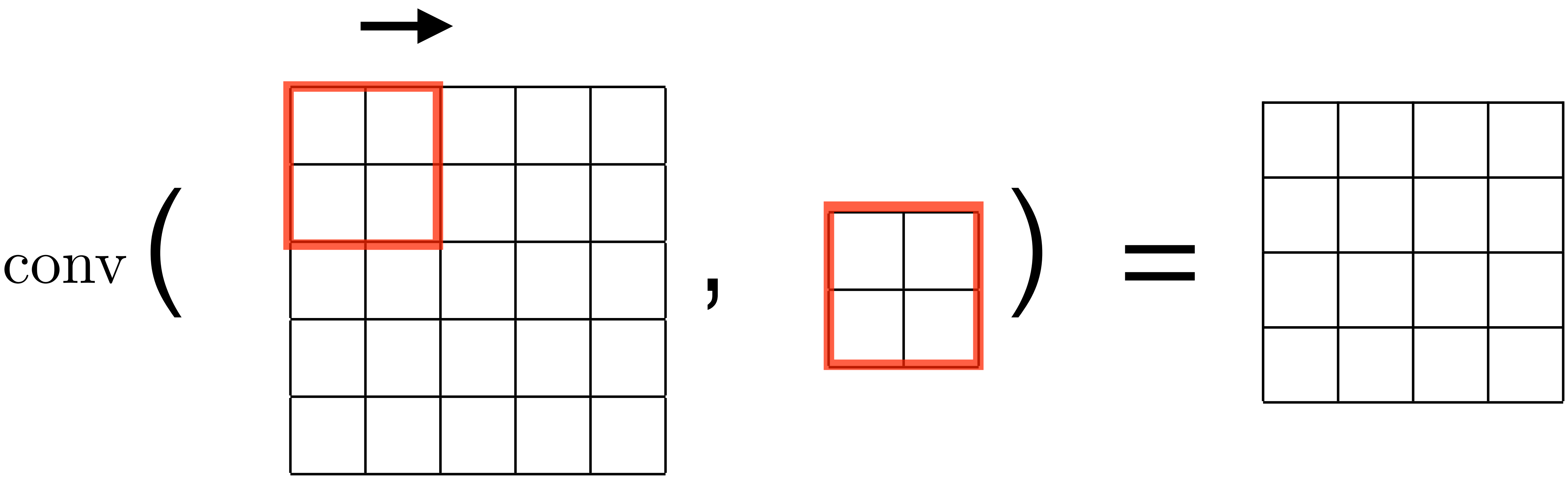


image (5x5)

kernel (2x2)

output (4x4)

Convolution layer properties - stride

stride = 3

kernel moves by 3 pixels

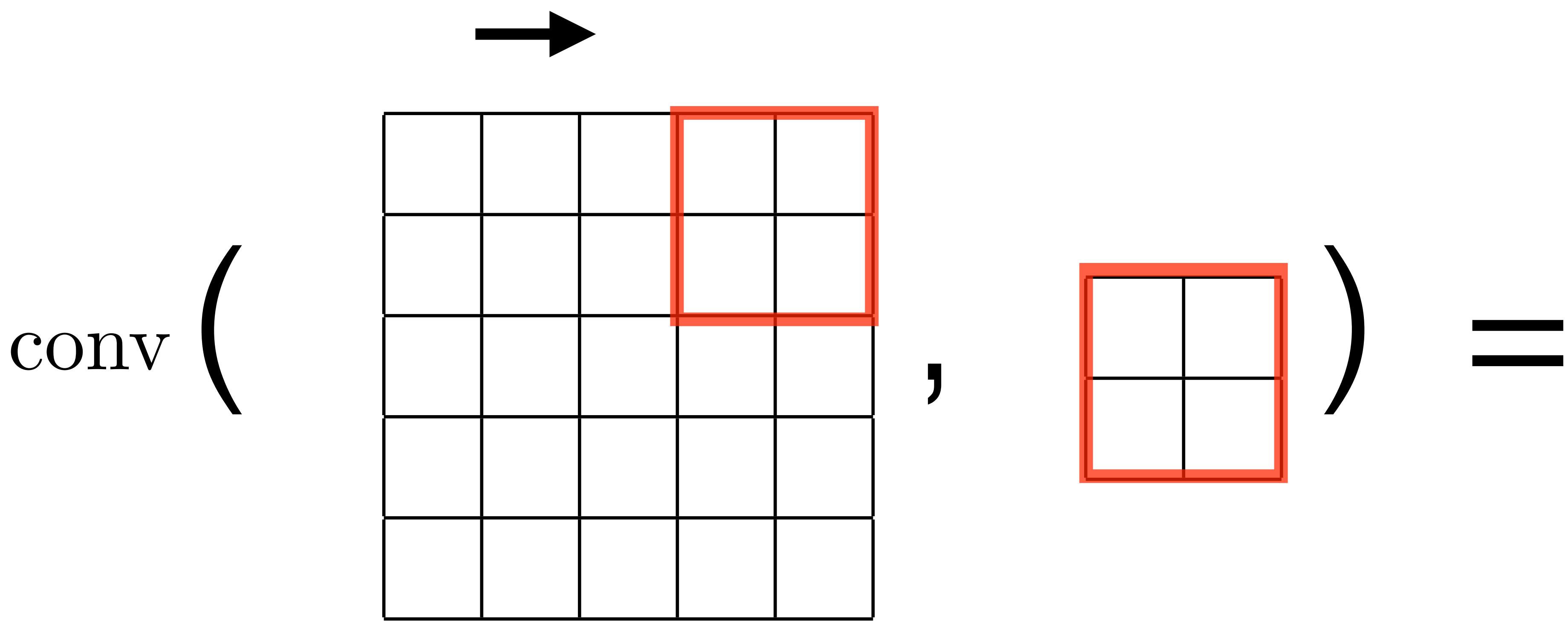


image
(5x5)

kernel
(2x2)

output
(**? x ?**)

Convolution layer properties - stride

stride = 3

kernel moves by 3 pixels

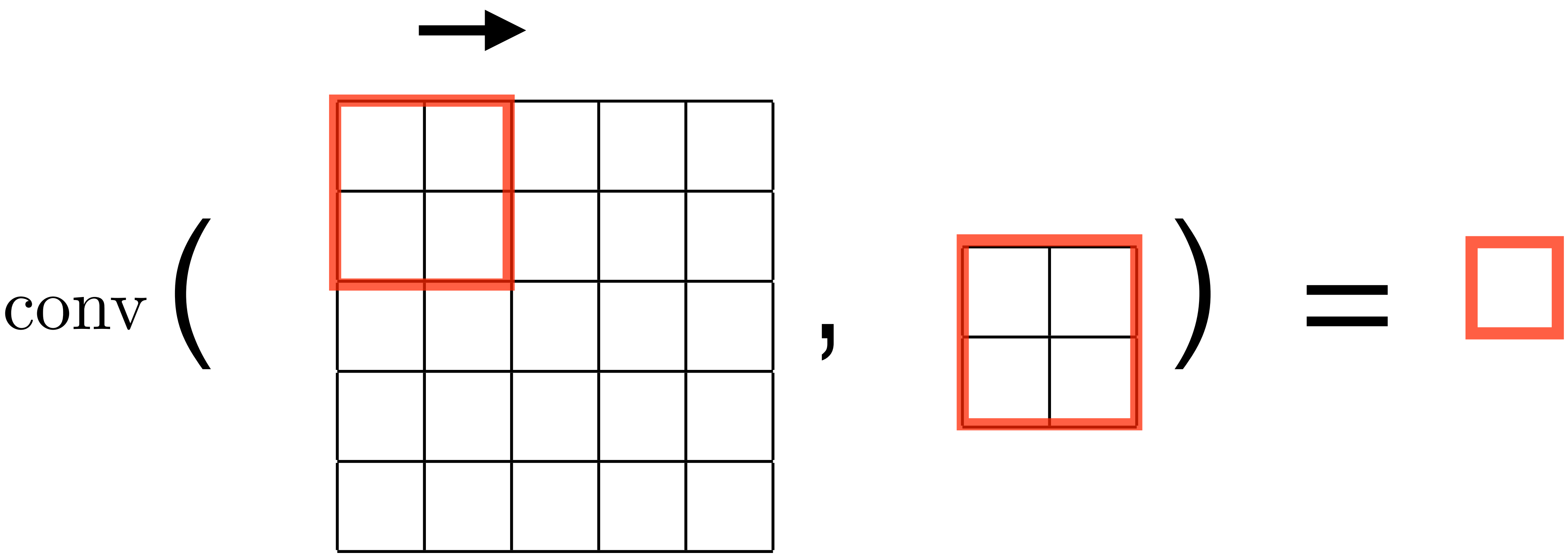


image
(5x5)

kernel
(2x2)

output
(**? x ?**)

Czech Technical University in Prague
Faculty of Electrical Engineering, Department of Cybernetics

Convolution layer properties - stride

stride = 3

kernel moves by 3 pixels

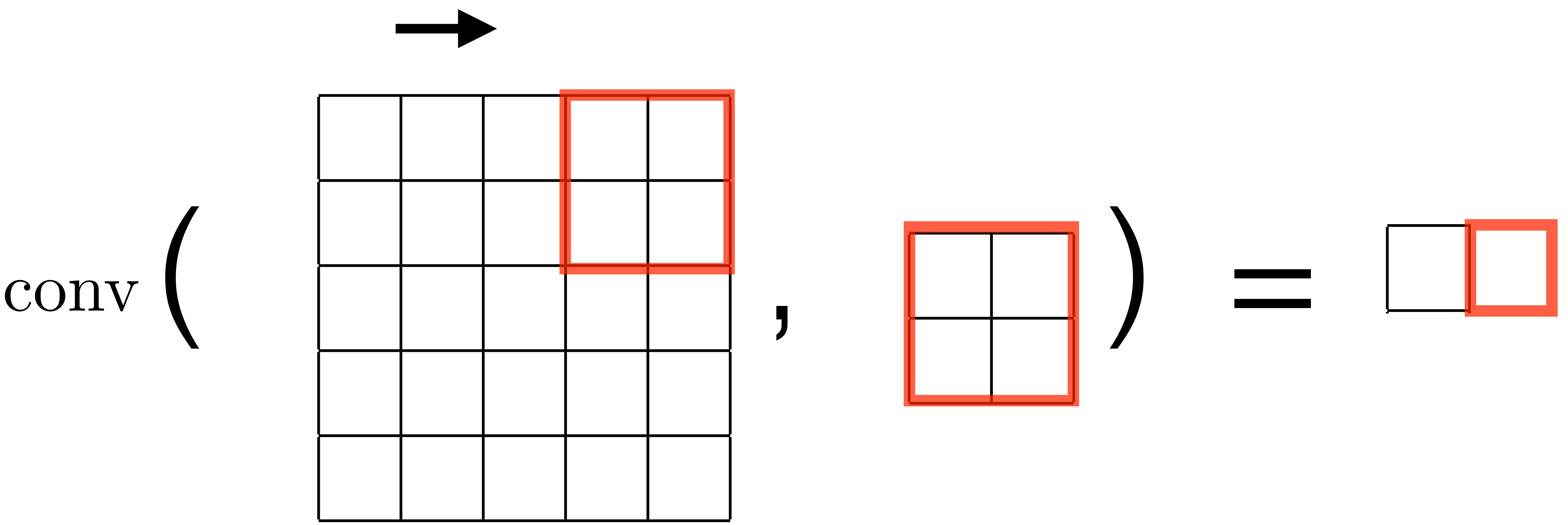


image (5x5)

kernel (2x2)

output (**? x ?**)

Convolution layer properties - stride

stride = 3

kernel moves by 3 pixels

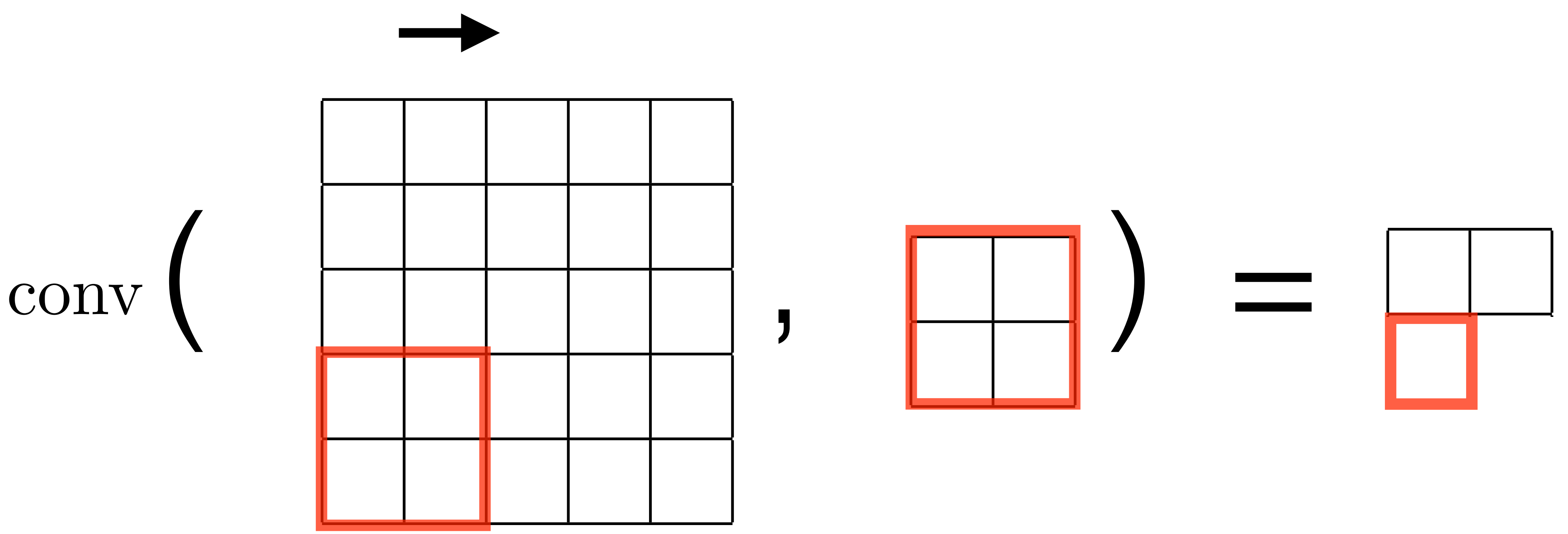


image (5x5) kernel (2x2) output (**? x ?**)

Convolution layer properties - stride

stride = 3

kernel moves by 3 pixels

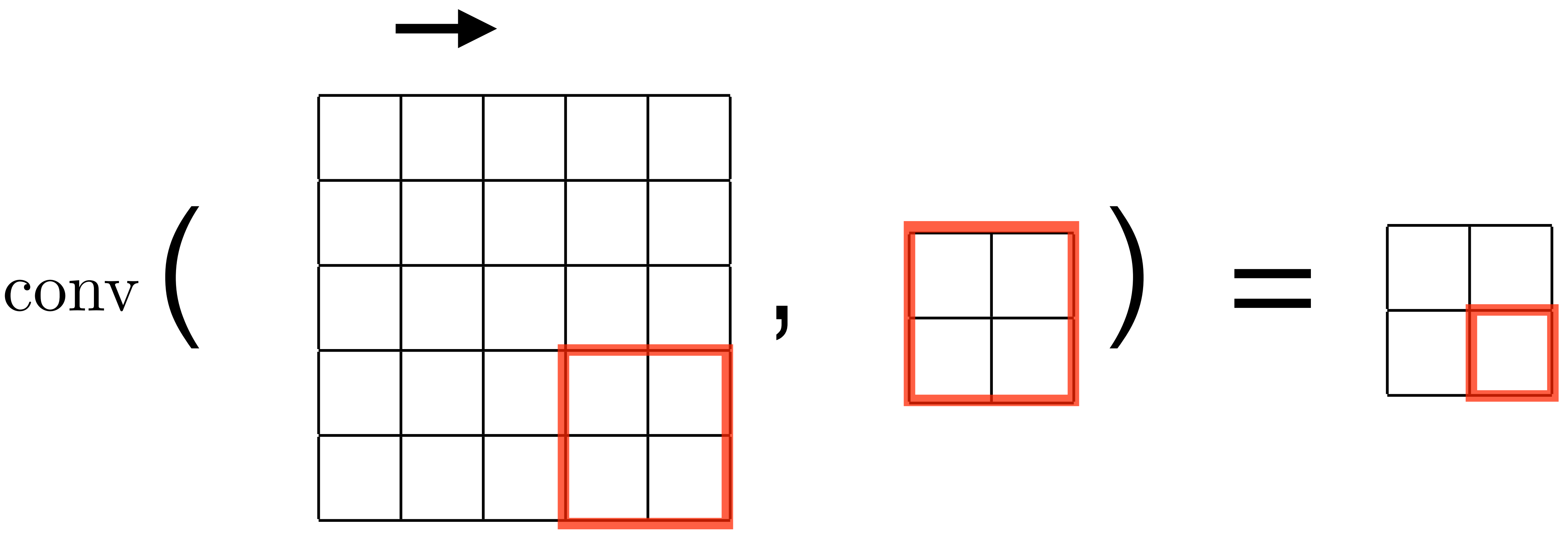


image (5x5)

kernel (2x2)

output (2x2)

Convolution layer properties - stride

$$M = (N-K) / \text{stride} + 1$$

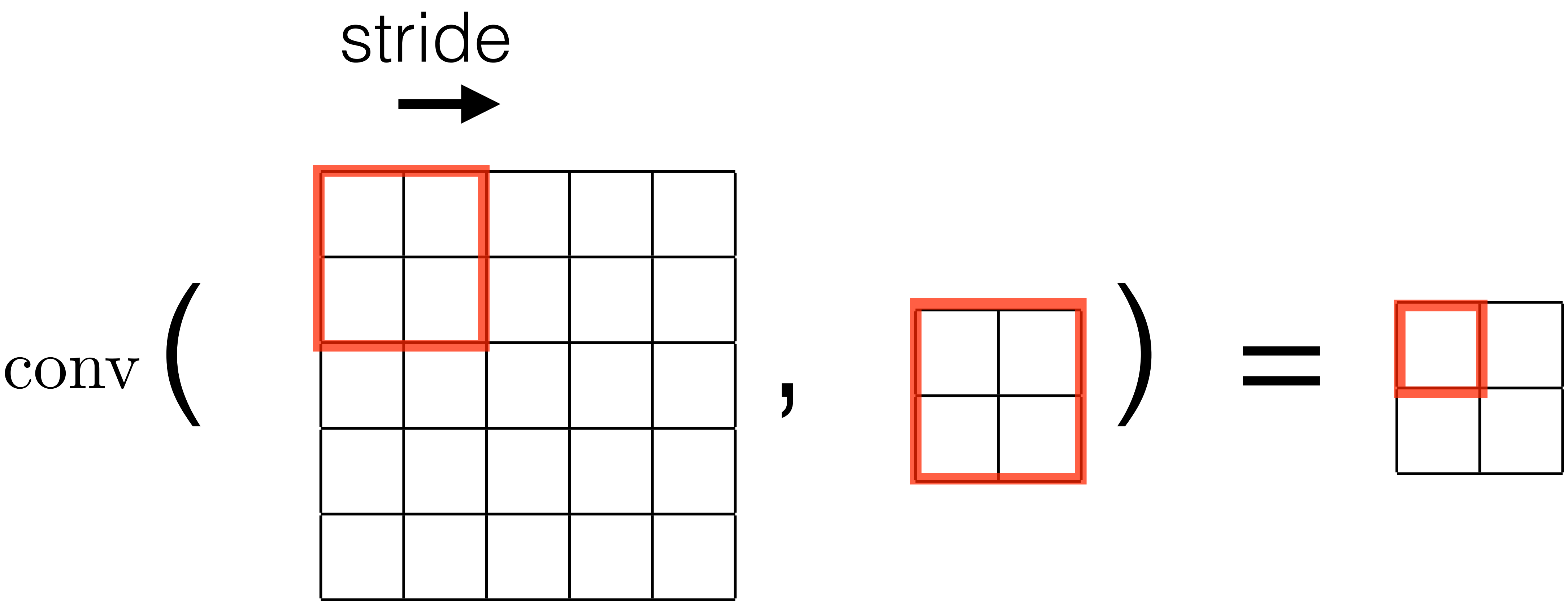


image (NxN)

kernel (KxK)

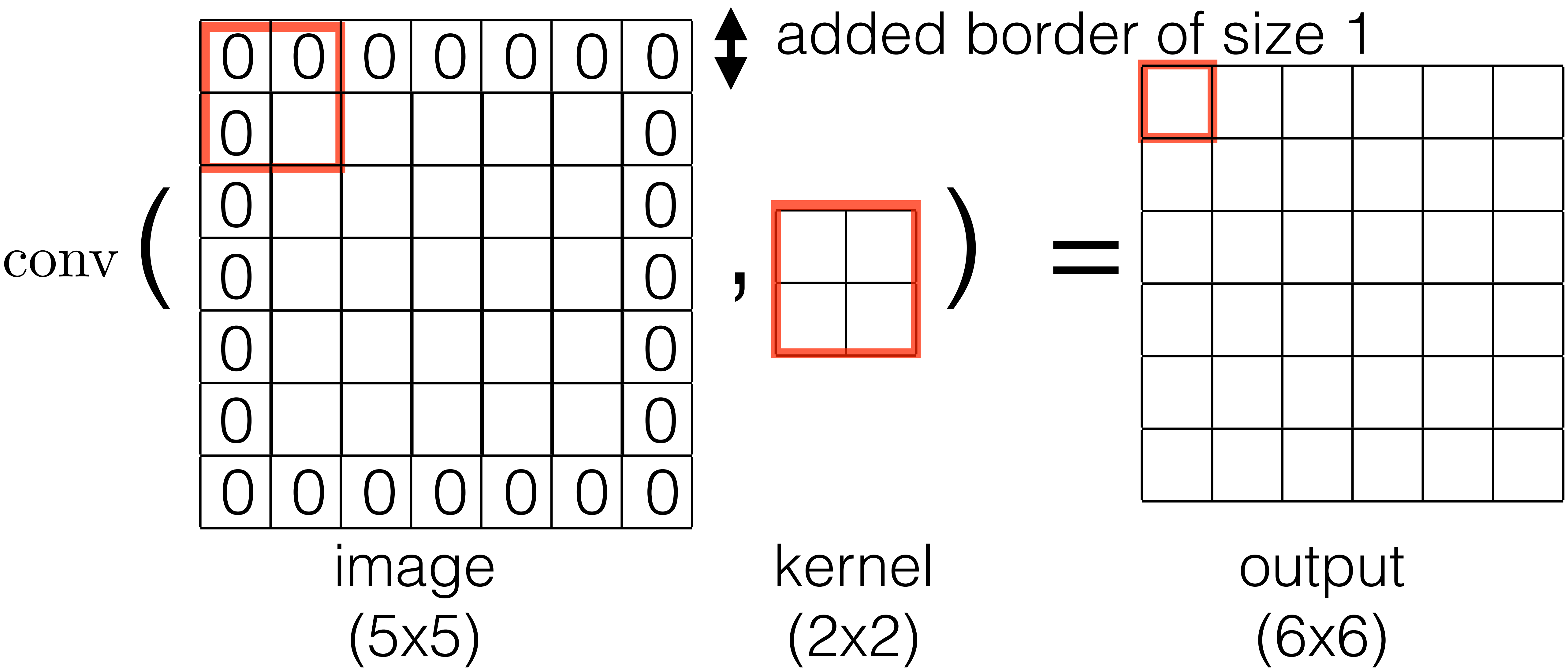
output (MxM)

e.g. $M = (5-2) / 3 + 1 = 2$

Czech Technical University in Prague
Faculty of Electrical Engineering, Department of Cybernetics

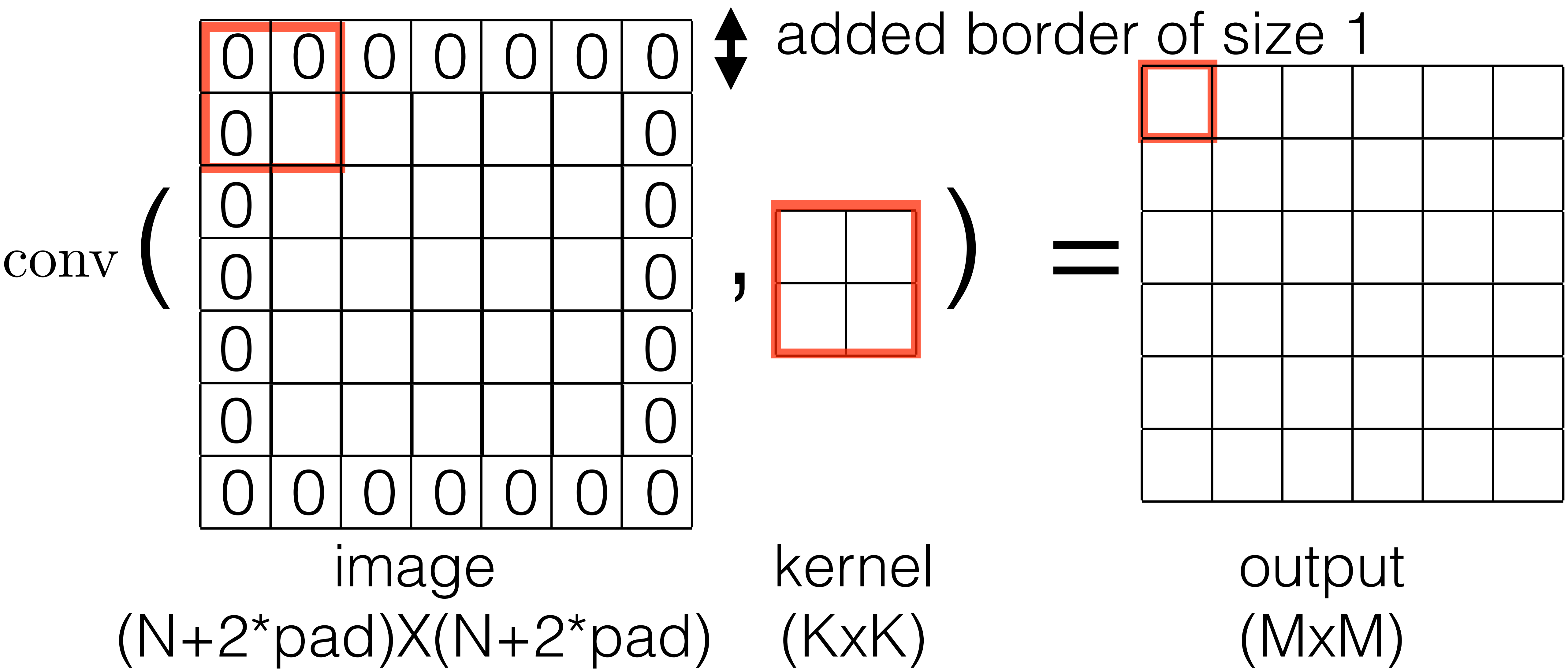
Convolution layer properties - pad

pad = 1



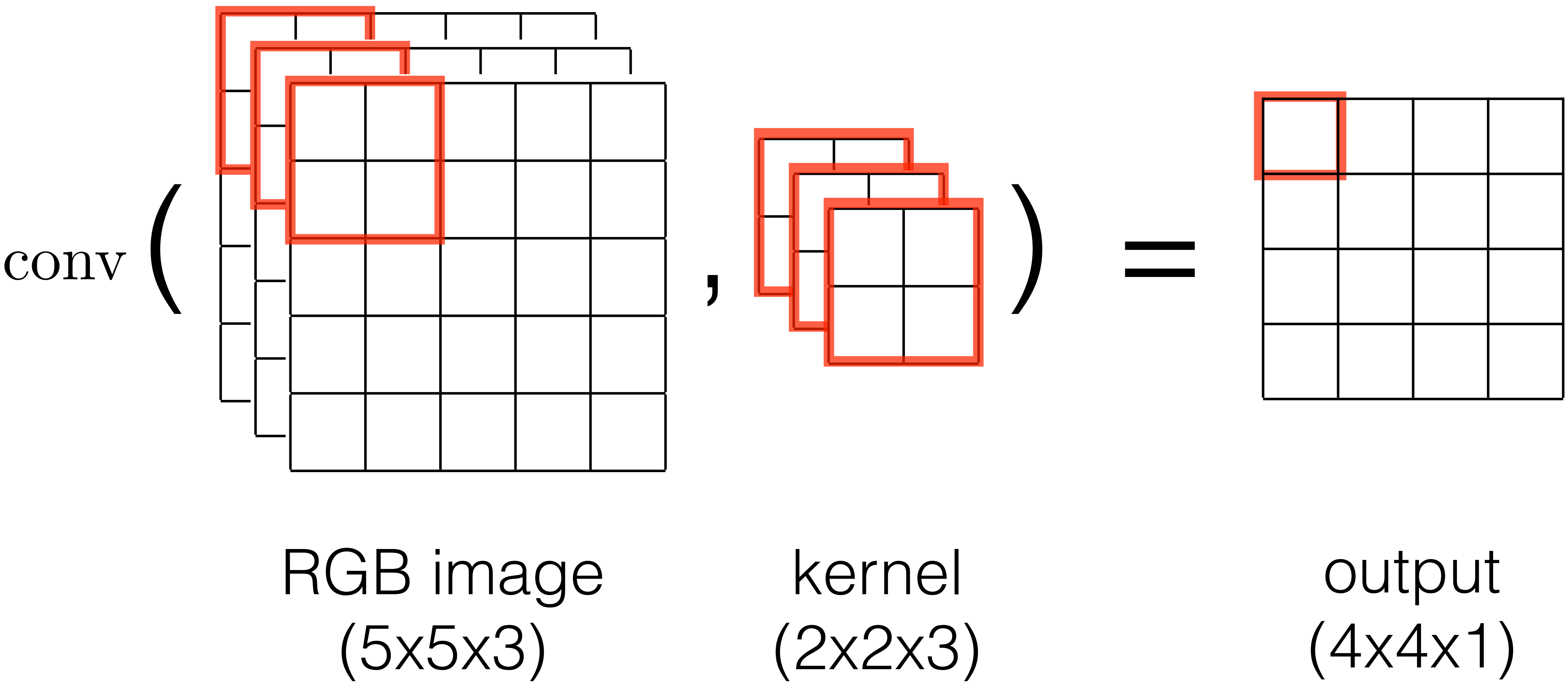
Convolution layer properties - pad

M = (N+2*pad-K) / stride + 1



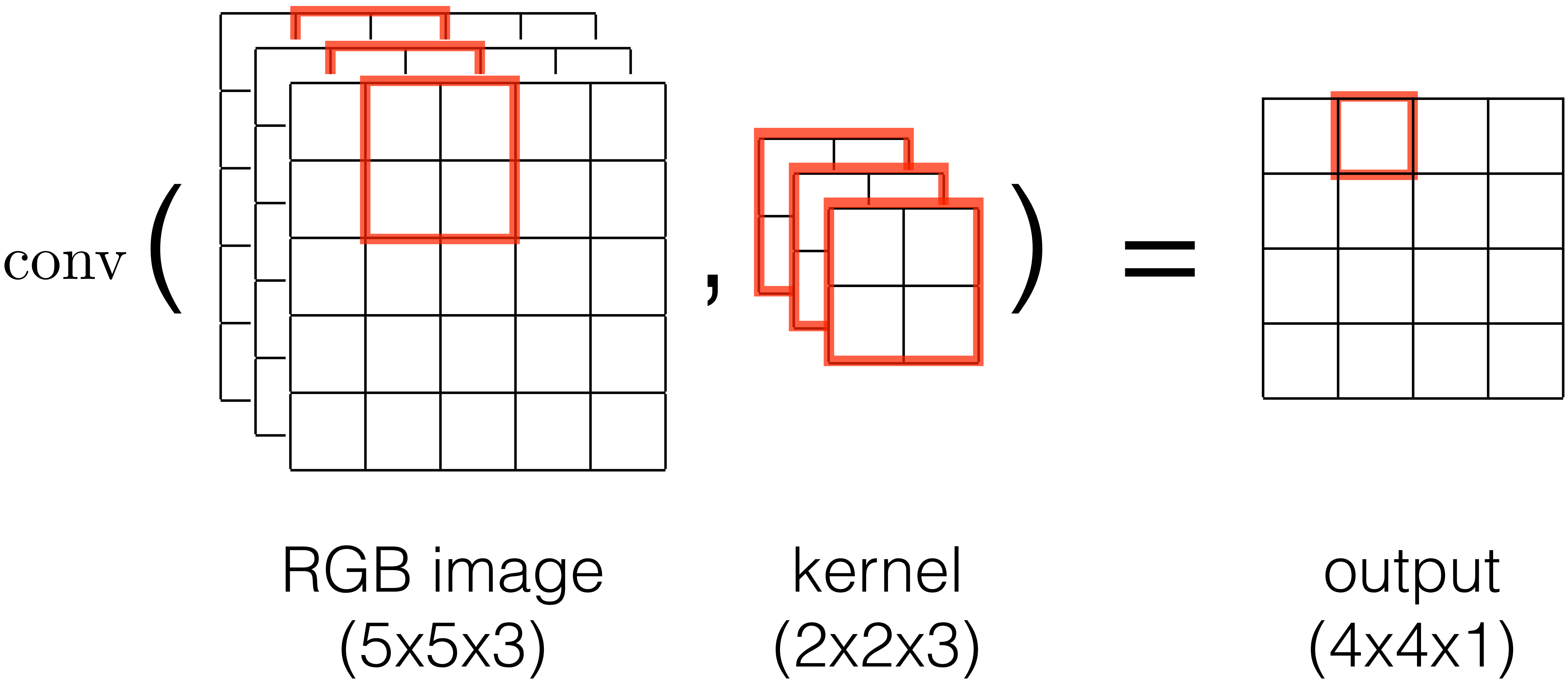
Czech Technical University in Prague
Faculty of Electrical Engineering, Department of Cybernetics

Multi-channel convolution

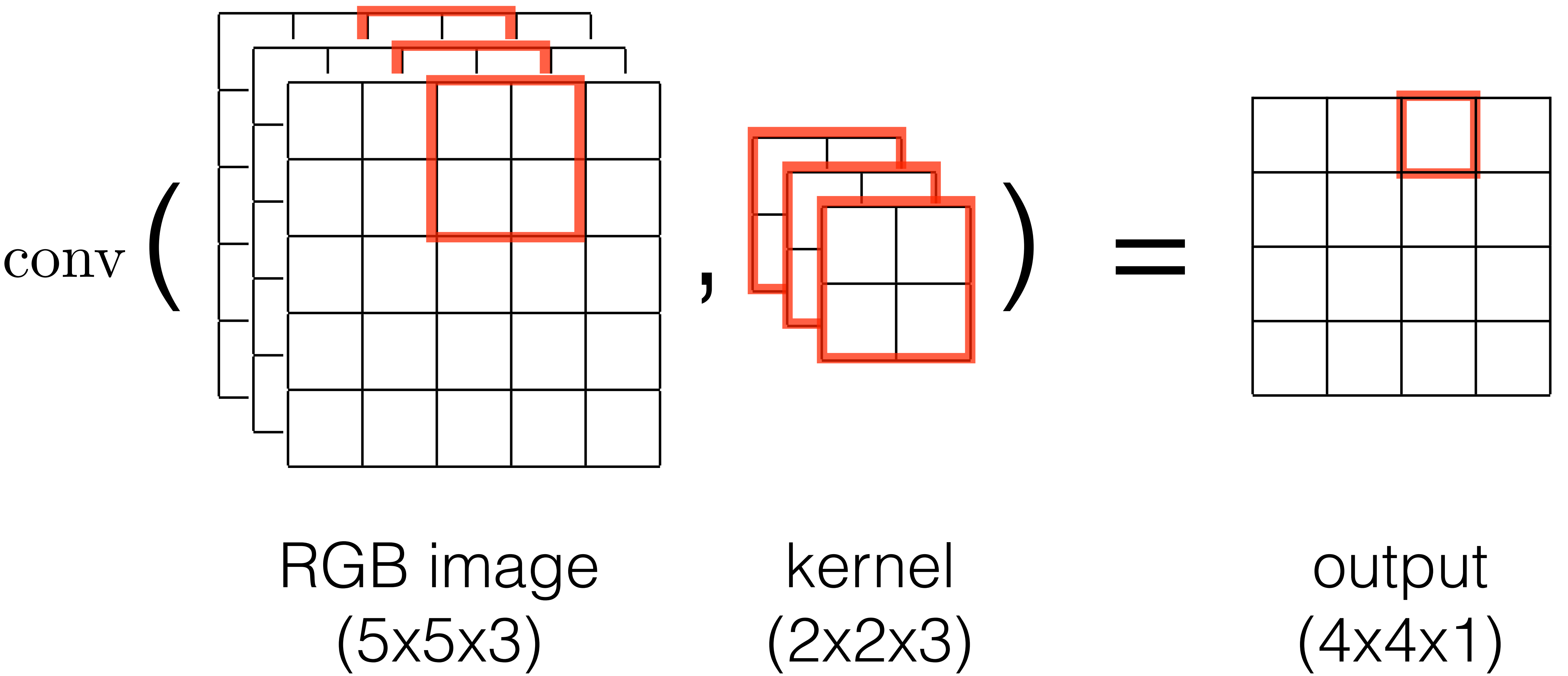


Czech Technical University in Prague
Faculty of Electrical Engineering, Department of Cybernetics

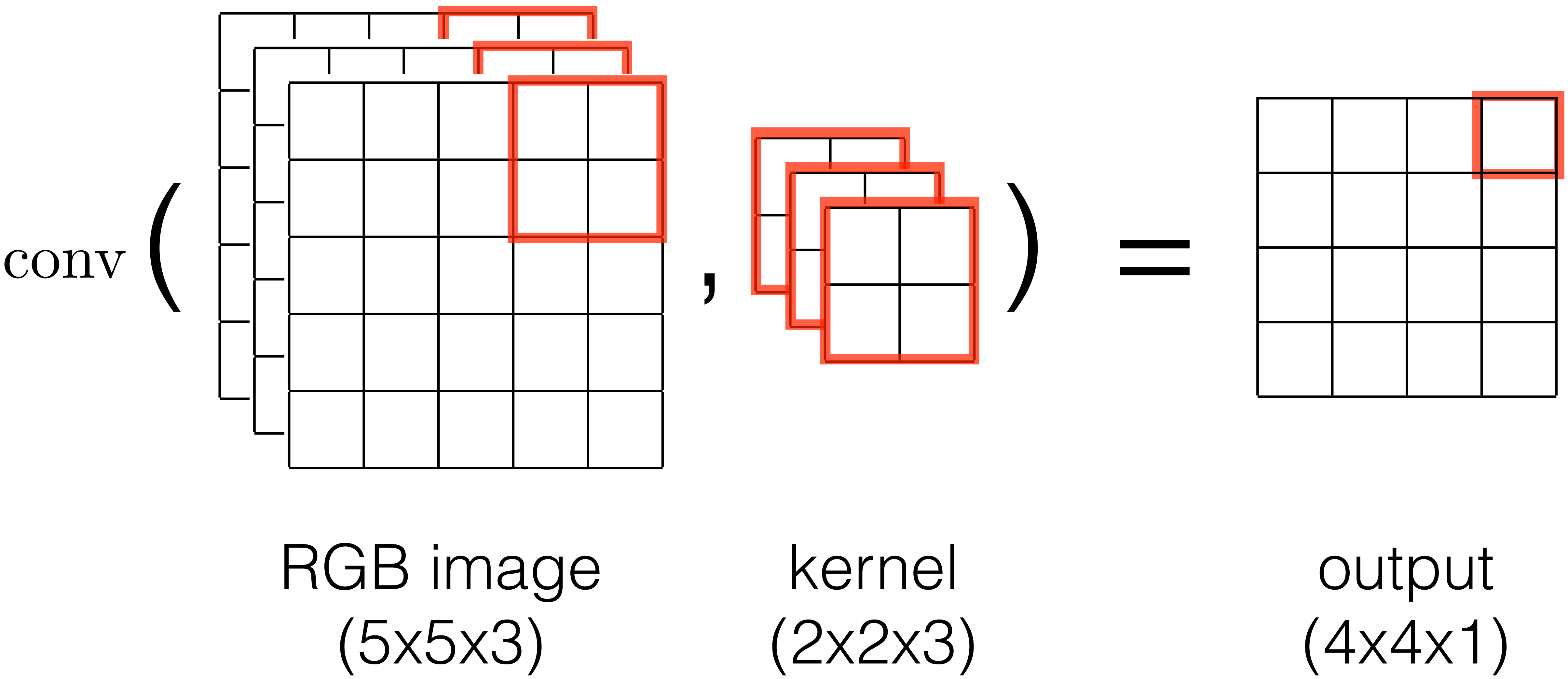
Multi-channel convolution



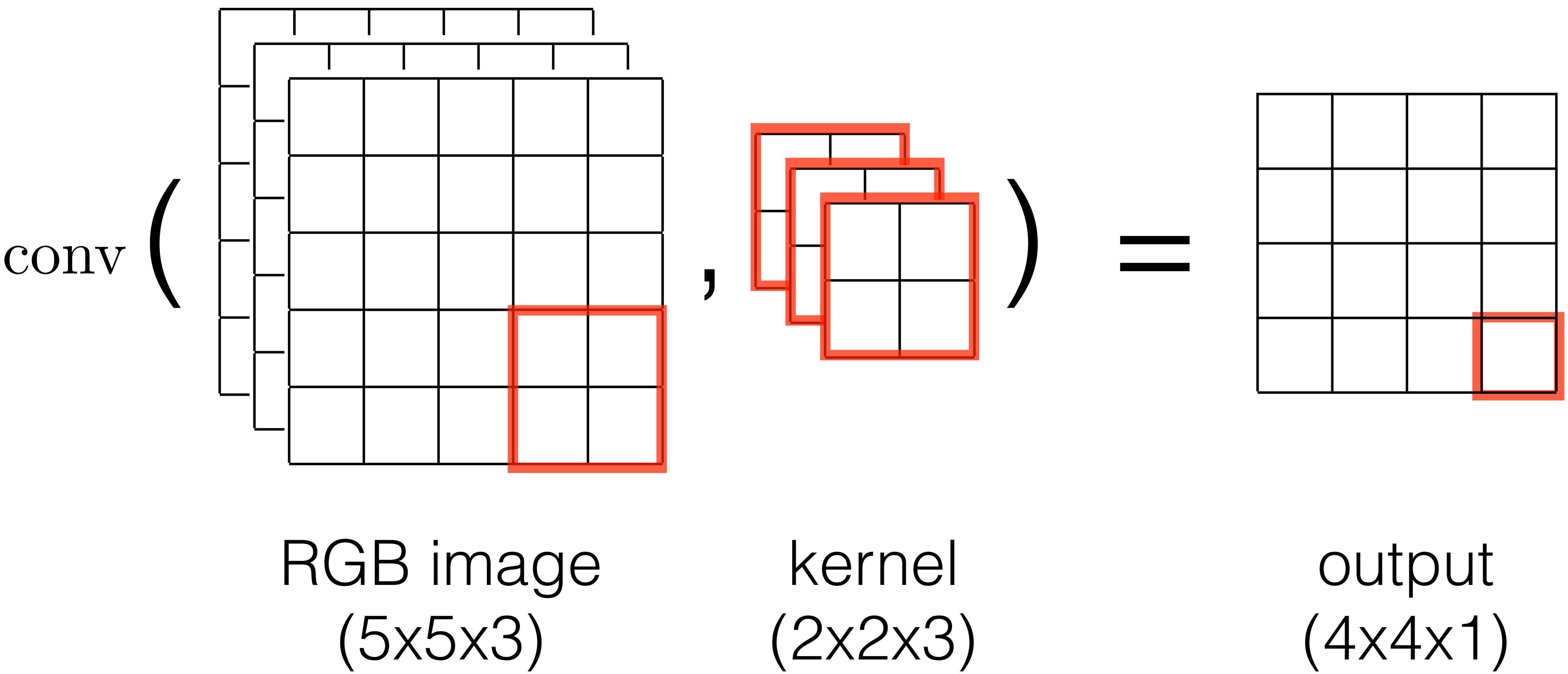
Multi-channel convolution



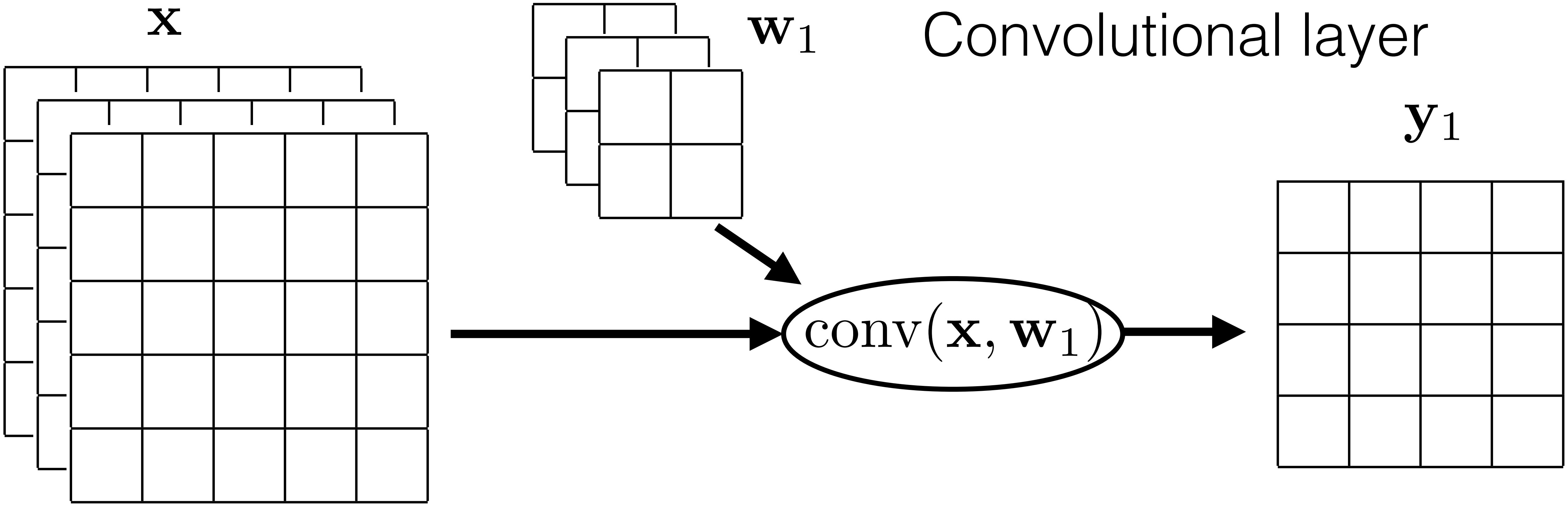
Multi-channel convolution



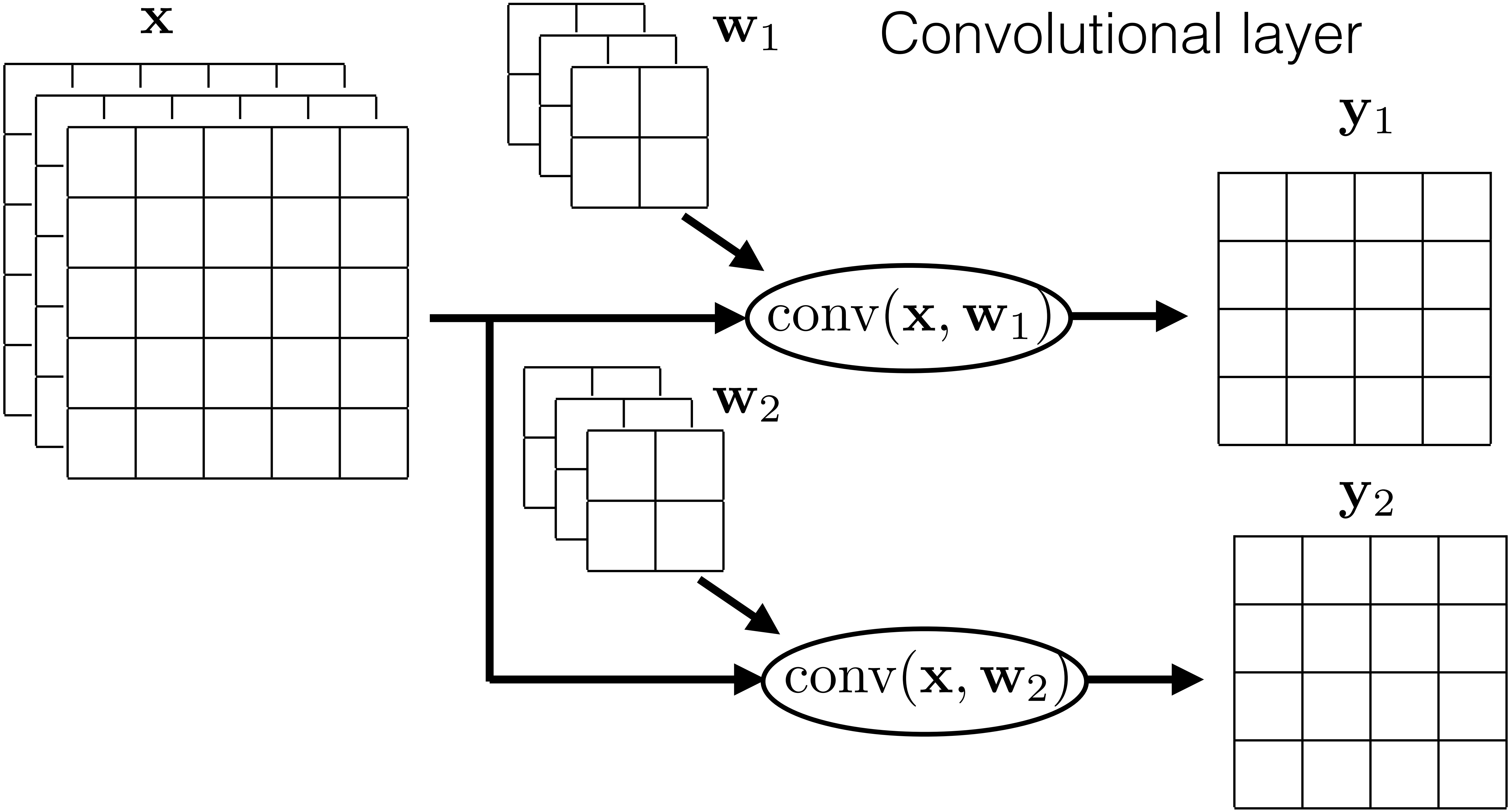
Multi-channel convolution



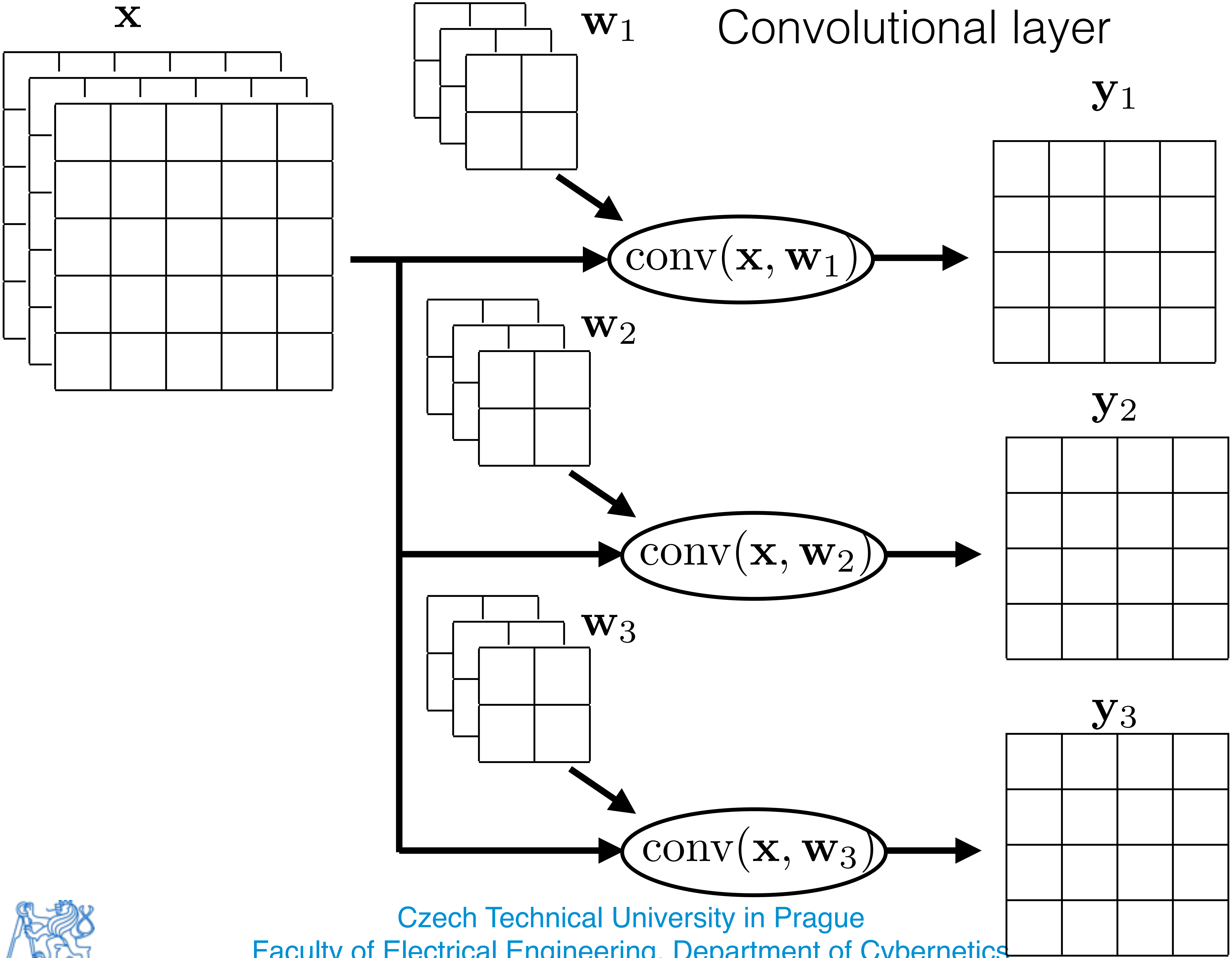
Convolutional layer



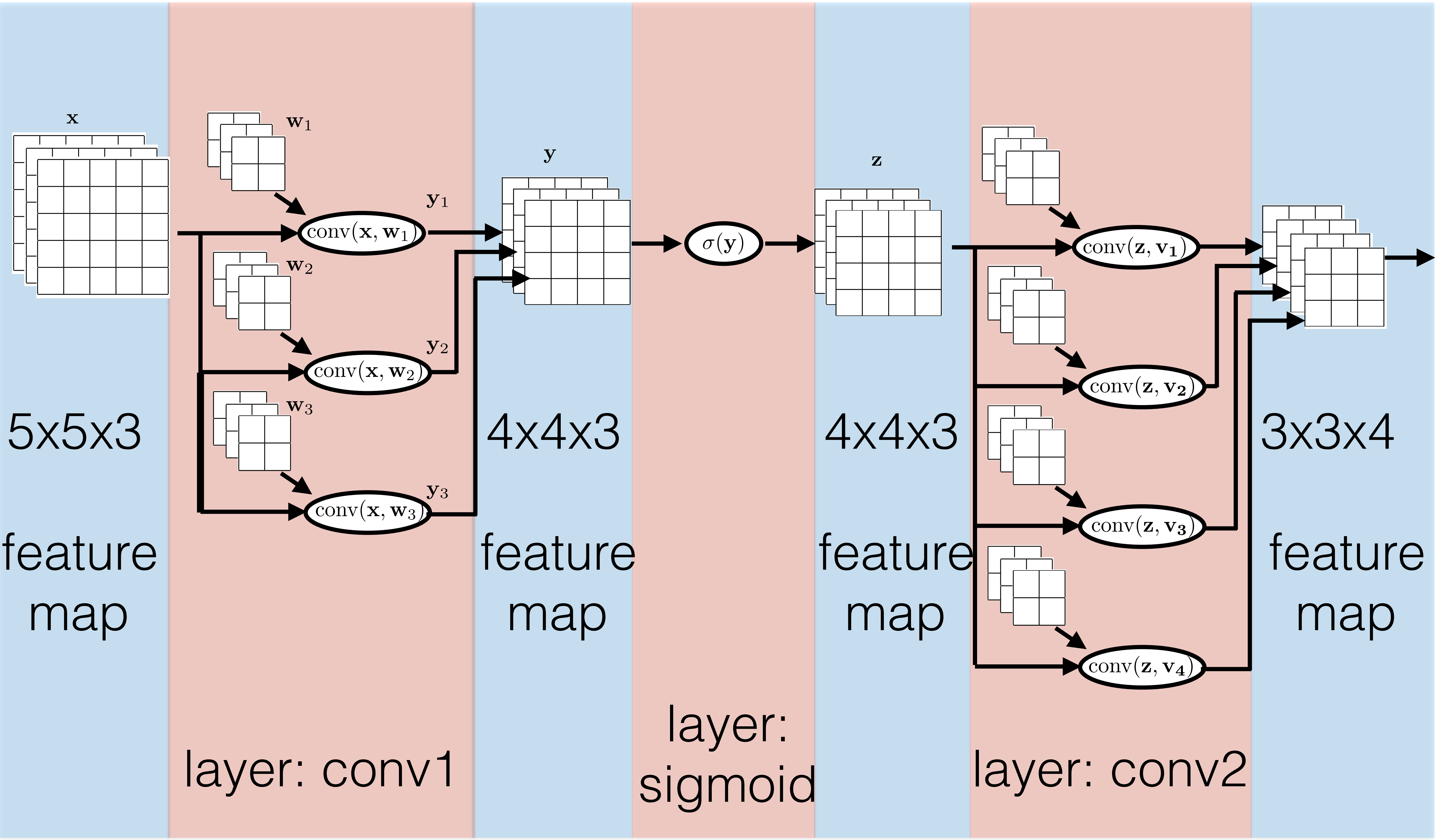
Convolutional layer



Convolutional layer

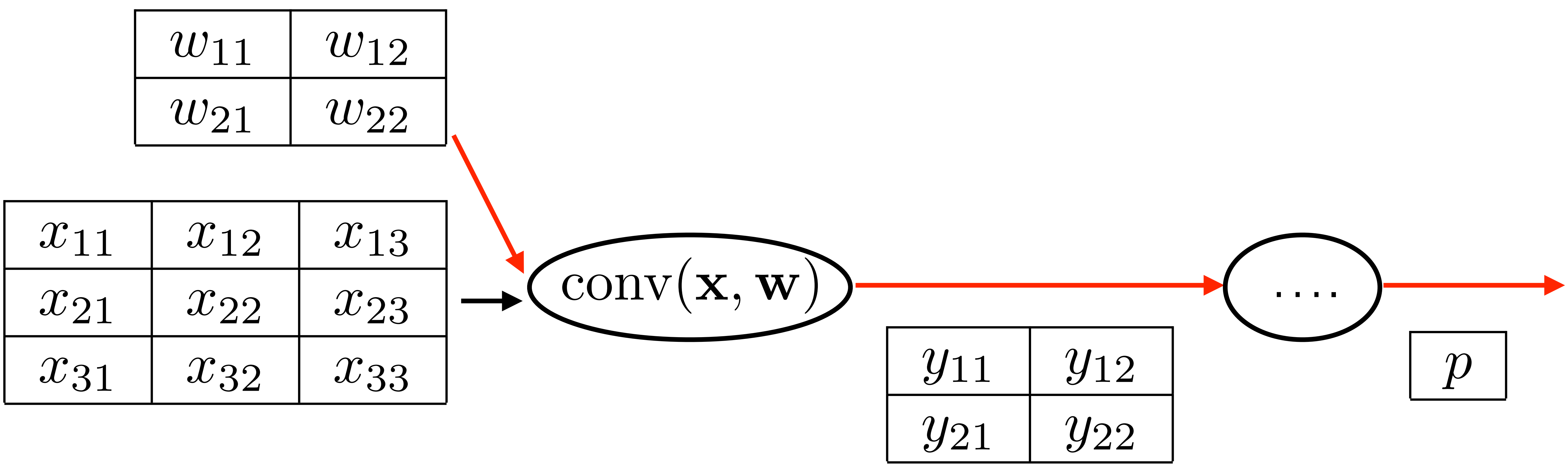


Convolutional network (ConvNet)



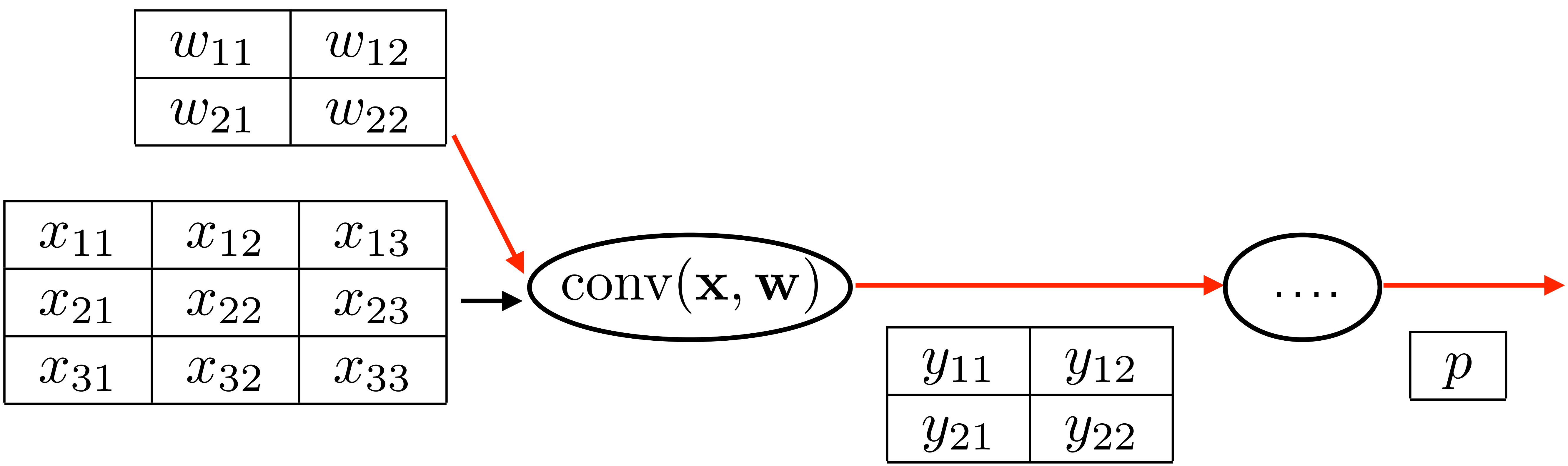
Convolution backward pass

Learning of convolutional neuron => backpropagation



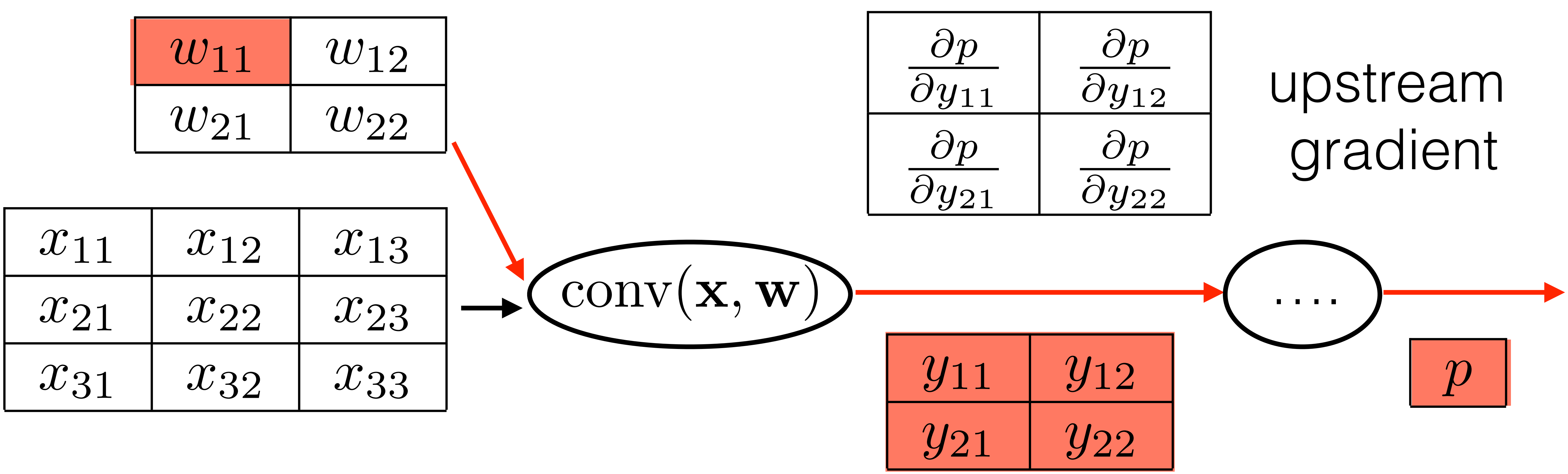
Convolution backward pass

$$\begin{array}{|c|c|}\hline \frac{\partial p}{\partial w_{11}} & \frac{\partial p}{\partial w_{12}} \\ \hline \frac{\partial p}{\partial w_{21}} & \frac{\partial p}{\partial w_{22}} \\ \hline \end{array} = ?$$



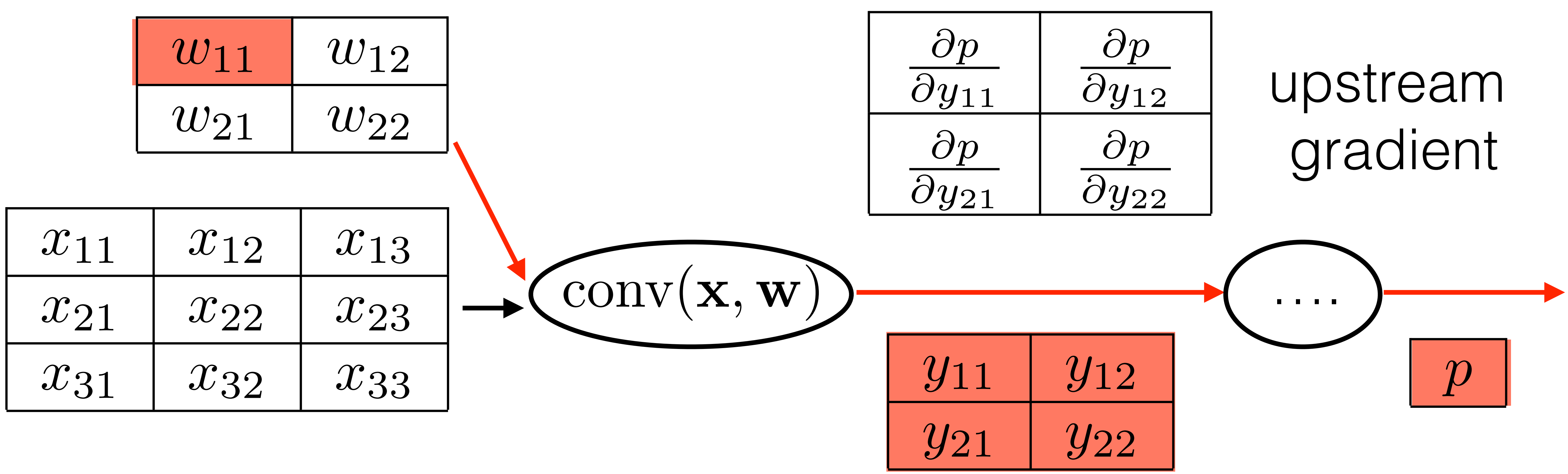
Convolution backward pass

$$\frac{\partial p}{\partial w_{11}} = \text{?}$$



Convolution backward pass

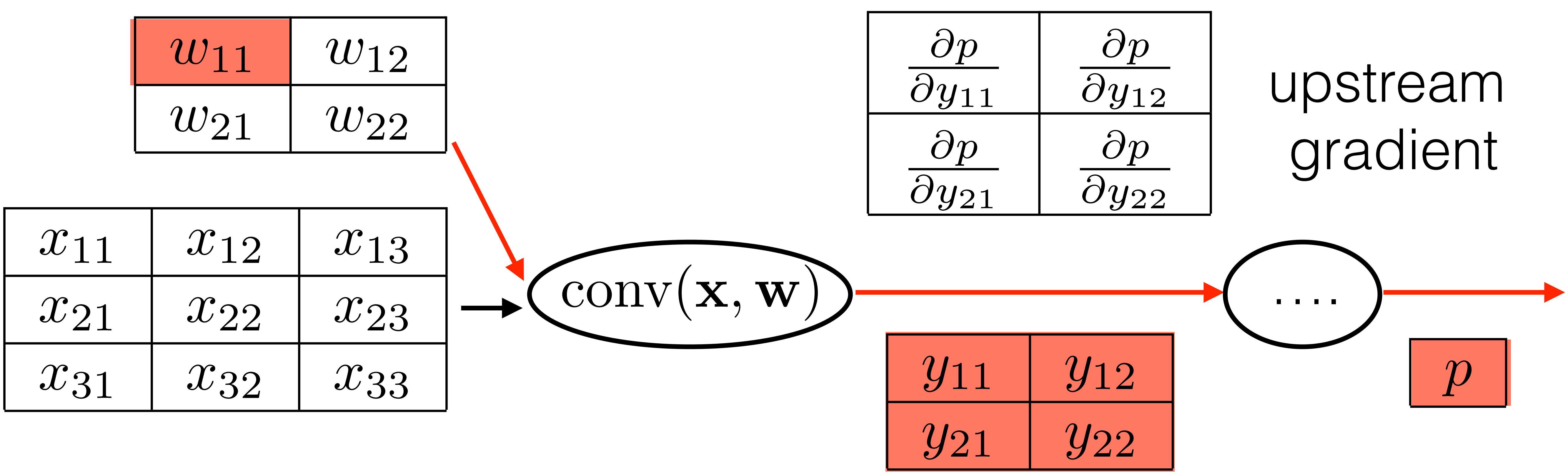
$$\frac{\partial p}{\partial w_{11}} = \frac{\partial p}{\partial y_{11}}\frac{\partial y_{11}}{\partial w_{11}} + \frac{\partial p}{\partial y_{12}}\frac{\partial y_{12}}{\partial w_{11}} + \frac{\partial p}{\partial y_{21}}\frac{\partial y_{21}}{\partial w_{11}} + \frac{\partial p}{\partial y_{22}}\frac{\partial y_{22}}{\partial w_{11}}$$



Convolution backward pass

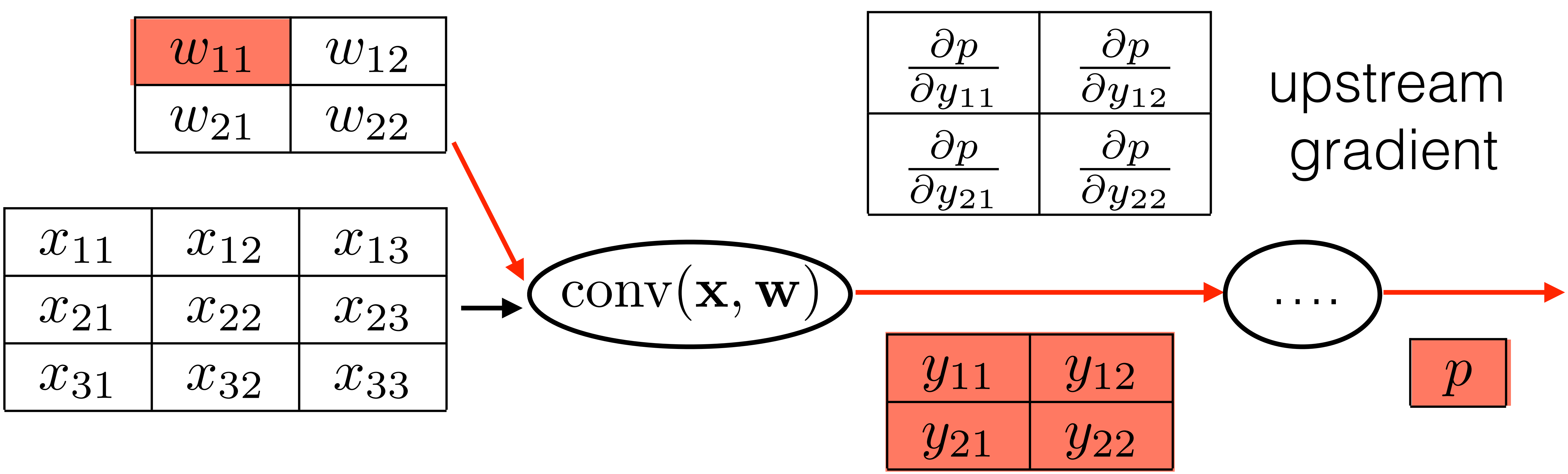
$$\frac{\partial p}{\partial w_{11}} = \frac{\partial p}{\partial y_{11}} \boxed{\frac{\partial y_{11}}{\partial w_{11}}} + \frac{\partial p}{\partial y_{12}} \frac{\partial y_{12}}{\partial w_{11}} + \frac{\partial p}{\partial y_{21}} \frac{\partial y_{21}}{\partial w_{11}} + \frac{\partial p}{\partial y_{22}} \frac{\partial y_{22}}{\partial w_{11}}$$

$$\frac{\partial y_{11}}{\partial w_{11}} = \frac{\partial (w_{11}x_{11} + w_{12}x_{12} + w_{21}x_{21} + w_{22}x_{22})}{\partial w_{11}} = x_{11}$$



Convolution backward pass

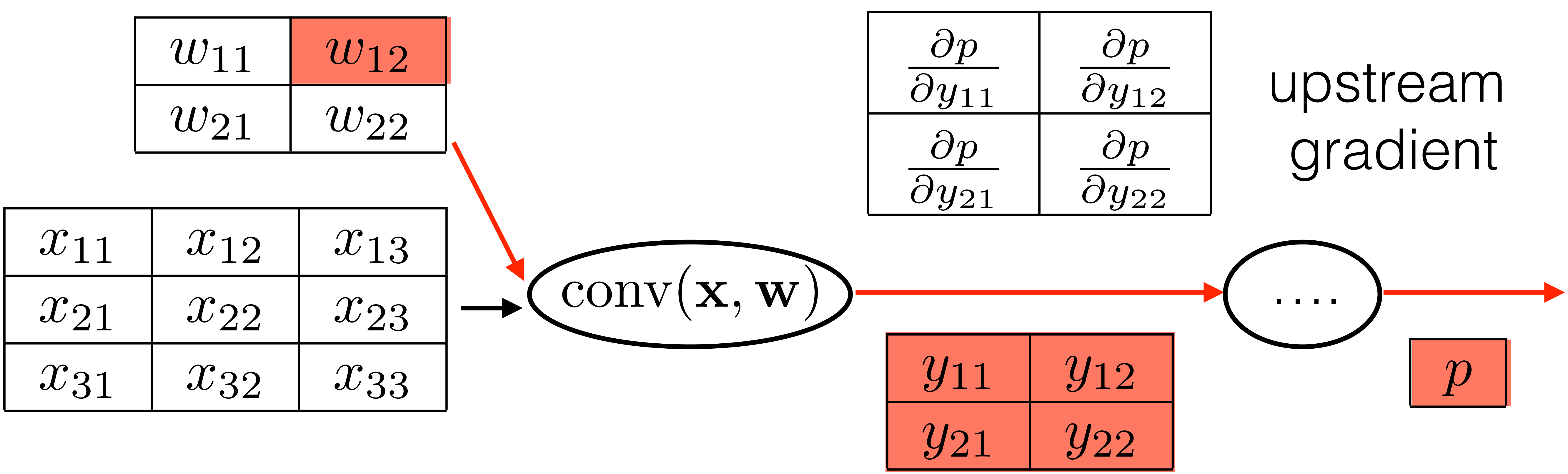
$$\frac{\partial p}{\partial w_{11}} = \frac{\partial p}{\partial y_{11}} x_{11} + \frac{\partial p}{\partial y_{12}} x_{12} + \frac{\partial p}{\partial y_{21}} x_{21} + \frac{\partial p}{\partial y_{22}} x_{22}$$



Convolution backward pass

$$\frac{\partial p}{\partial w_{11}} = \frac{\partial p}{\partial y_{11}} x_{11} + \frac{\partial p}{\partial y_{12}} x_{12} + \frac{\partial p}{\partial y_{21}} x_{21} + \frac{\partial p}{\partial y_{22}} x_{22}$$

$$\frac{\partial p}{\partial w_{12}} = \frac{\partial p}{\partial y_{11}} x_{12} + \frac{\partial p}{\partial y_{12}} x_{13} + \frac{\partial p}{\partial y_{21}} x_{22} + \frac{\partial p}{\partial y_{22}} x_{23}$$

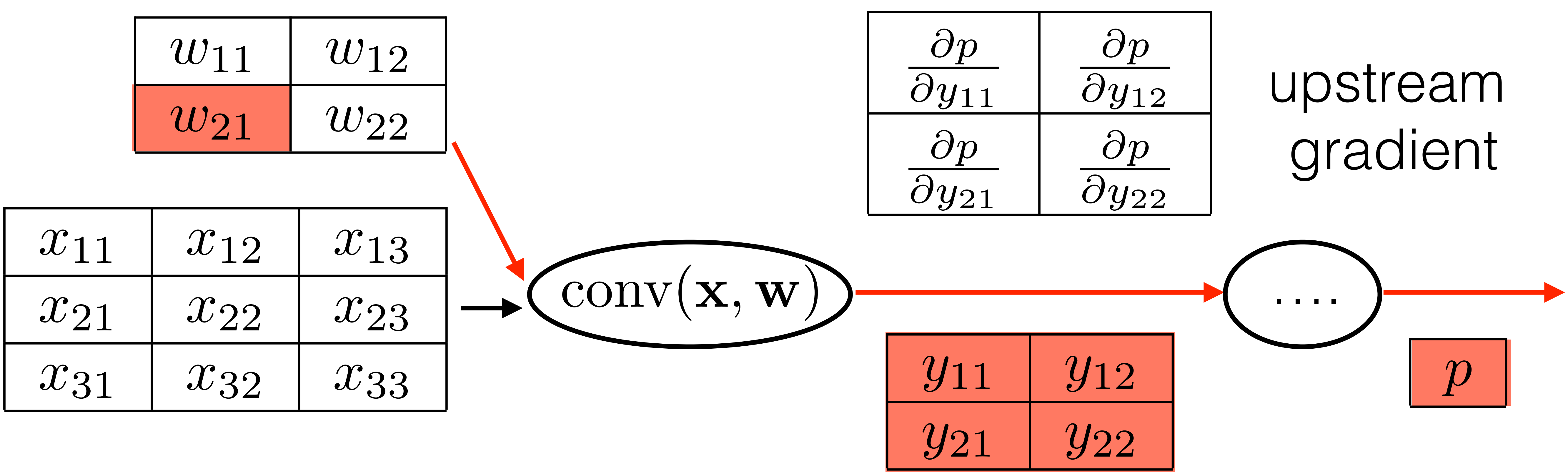


Convolution backward pass

$$\frac{\partial p}{\partial w_{11}} = \frac{\partial p}{\partial y_{11}} x_{11} + \frac{\partial p}{\partial y_{12}} x_{12} + \frac{\partial p}{\partial y_{21}} x_{21} + \frac{\partial p}{\partial y_{22}} x_{22}$$

$$\frac{\partial p}{\partial w_{12}} = \frac{\partial p}{\partial y_{11}} x_{12} + \frac{\partial p}{\partial y_{12}} x_{13} + \frac{\partial p}{\partial y_{21}} x_{22} + \frac{\partial p}{\partial y_{22}} x_{23}$$

$$\frac{\partial p}{\partial w_{21}} = \frac{\partial p}{\partial y_{11}} x_{21} + \frac{\partial p}{\partial y_{12}} x_{22} + \frac{\partial p}{\partial y_{21}} x_{31} + \frac{\partial p}{\partial y_{22}} x_{32}$$



Convolution backward pass

$$\frac{\partial p}{\partial w_{11}} = \frac{\partial p}{\partial y_{11}} x_{11} + \frac{\partial p}{\partial y_{12}} x_{12} + \frac{\partial p}{\partial y_{21}} x_{21} + \frac{\partial p}{\partial y_{22}} x_{22}$$

$$\frac{\partial p}{\partial w_{12}} = \frac{\partial p}{\partial y_{11}} x_{12} + \frac{\partial p}{\partial y_{12}} x_{13} + \frac{\partial p}{\partial y_{21}} x_{22} + \frac{\partial p}{\partial y_{22}} x_{23}$$

$$\frac{\partial p}{\partial w_{21}} = \frac{\partial p}{\partial y_{11}} x_{21} + \frac{\partial p}{\partial y_{12}} x_{22} + \frac{\partial p}{\partial y_{21}} x_{31} + \frac{\partial p}{\partial y_{22}} x_{32}$$

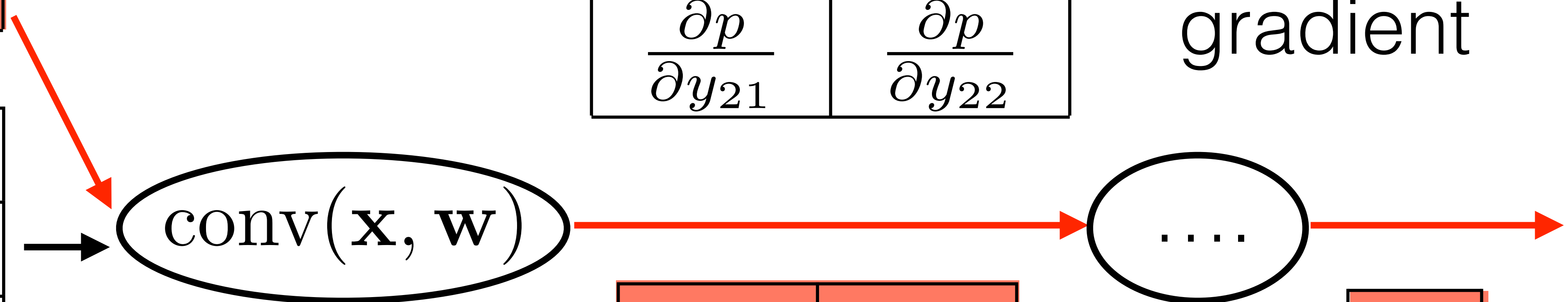
$$\frac{\partial p}{\partial w_{22}} = \frac{\partial p}{\partial y_{11}} x_{22} + \frac{\partial p}{\partial y_{12}} x_{23} + \frac{\partial p}{\partial y_{21}} x_{32} + \frac{\partial p}{\partial y_{22}} x_{33}$$

w_{11}	w_{12}
w_{21}	w_{22}

$\frac{\partial p}{\partial y_{11}}$	$\frac{\partial p}{\partial y_{12}}$
$\frac{\partial p}{\partial y_{21}}$	$\frac{\partial p}{\partial y_{22}}$

upstream gradient

x_{11}	x_{12}	x_{13}
x_{21}	x_{22}	x_{23}
x_{31}	x_{32}	x_{33}



y_{11}	y_{12}
y_{21}	y_{22}

p

Czech Technical University in Prague
Faculty of Electrical Engineering, Department of Cybernetics

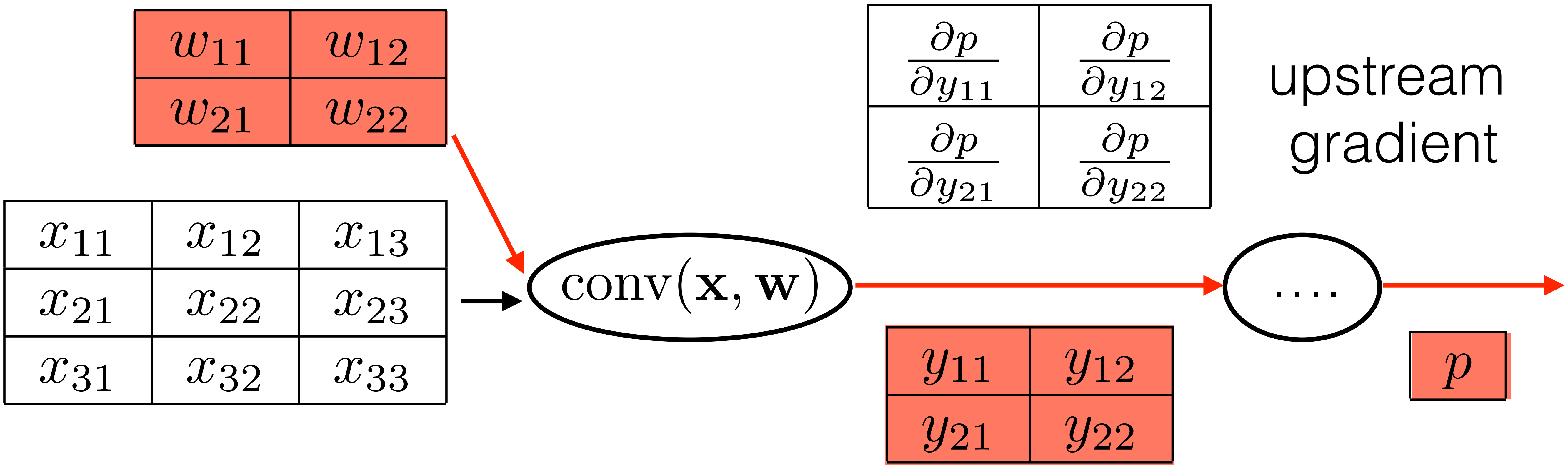
Convolution backward pass

$$\frac{\partial p}{\partial w_{11}} = \frac{\partial p}{\partial y_{11}} x_{11} + \frac{\partial p}{\partial y_{12}} x_{12} + \frac{\partial p}{\partial y_{21}} x_{21} + \frac{\partial p}{\partial y_{22}} x_{22}$$

$$\frac{\partial p}{\partial w_{12}} = \frac{\partial p}{\partial y_{11}} x_{12} + \frac{\partial p}{\partial y_{12}} x_{13} + \frac{\partial p}{\partial y_{21}} x_{22} + \frac{\partial p}{\partial y_{22}} x_{23}$$

$$\frac{\partial p}{\partial w_{21}} = \frac{\partial p}{\partial y_{11}} x_{21} + \frac{\partial p}{\partial y_{12}} x_{22} + \frac{\partial p}{\partial y_{21}} x_{31} + \frac{\partial p}{\partial y_{22}} x_{32}$$

$$\frac{\partial p}{\partial w_{22}} = \frac{\partial p}{\partial y_{11}} x_{22} + \frac{\partial p}{\partial y_{12}} x_{23} + \frac{\partial p}{\partial y_{21}} x_{32} + \frac{\partial p}{\partial y_{22}} x_{33}$$



Convolution backward pass

$$\frac{\partial p}{\partial w_{11}} = \frac{\partial p}{\partial y_{11}}x_{11} + \frac{\partial p}{\partial y_{12}}x_{12} + \frac{\partial p}{\partial y_{21}}x_{21} + \frac{\partial p}{\partial y_{22}}x_{22}$$

$$\frac{\partial p}{\partial w_{12}} = \frac{\partial p}{\partial y_{11}}x_{12} + \frac{\partial p}{\partial y_{12}}x_{13} + \frac{\partial p}{\partial y_{21}}x_{22} + \frac{\partial p}{\partial y_{22}}x_{23}$$

$$\frac{\partial p}{\partial w_{21}} = \frac{\partial p}{\partial y_{11}}x_{21} + \frac{\partial p}{\partial y_{12}}x_{22} + \frac{\partial p}{\partial y_{21}}x_{31} + \frac{\partial p}{\partial y_{22}}x_{32}$$

$$\frac{\partial p}{\partial w_{22}} = \frac{\partial p}{\partial y_{11}}x_{22} + \frac{\partial p}{\partial y_{12}}x_{23} + \frac{\partial p}{\partial y_{21}}x_{32} + \frac{\partial p}{\partial y_{22}}x_{33}$$

$$\begin{array}{|c|c|}\hline \frac{\partial p}{\partial w_{11}} & \frac{\partial p}{\partial w_{12}} \\ \hline \frac{\partial p}{\partial w_{21}} & \frac{\partial p}{\partial w_{22}} \\ \hline \end{array} = \mathrm{conv}\left(\begin{array}{|c|c|c|}\hline x_{11} & x_{12} & x_{13} \\ \hline x_{21} & x_{22} & x_{23} \\ \hline x_{31} & x_{32} & x_{33} \\ \hline \end{array}, \begin{array}{|c|c|}\hline \frac{\partial p}{\partial y_{11}} & \frac{\partial p}{\partial y_{12}} \\ \hline \frac{\partial p}{\partial y_{21}} & \frac{\partial p}{\partial y_{22}} \\ \hline \end{array} \right)$$

Convolution backward pass

$$\frac{\partial p}{\partial w_{11}} = \frac{\partial p}{\partial y_{11}} x_{11} + \frac{\partial p}{\partial y_{12}} x_{12} + \frac{\partial p}{\partial y_{21}} x_{21} + \frac{\partial p}{\partial y_{22}} x_{22}$$

$$\frac{\partial p}{\partial w_{12}} = \frac{\partial p}{\partial y_{11}} x_{12} + \frac{\partial p}{\partial y_{12}} x_{13} + \frac{\partial p}{\partial y_{21}} x_{22} + \frac{\partial p}{\partial y_{22}} x_{23}$$

$$\frac{\partial p}{\partial w_{21}} = \frac{\partial p}{\partial y_{11}} x_{21} + \frac{\partial p}{\partial y_{12}} x_{22} + \frac{\partial p}{\partial y_{21}} x_{31} + \frac{\partial p}{\partial y_{22}} x_{32}$$

$$\frac{\partial p}{\partial w_{22}} = \frac{\partial p}{\partial y_{11}} x_{22} + \frac{\partial p}{\partial y_{12}} x_{23} + \frac{\partial p}{\partial y_{21}} x_{32} + \frac{\partial p}{\partial y_{22}} x_{33}$$

$$\begin{array}{|c|c|}\hline \frac{\partial p}{\partial w_{11}} & \frac{\partial p}{\partial w_{12}} \\ \hline \frac{\partial p}{\partial w_{21}} & \frac{\partial p}{\partial w_{22}} \\ \hline \end{array} = \mathrm{conv}\left(\begin{array}{|c|c|c|}\hline x_{11} & x_{12} & x_{13} \\ \hline x_{21} & x_{22} & x_{23} \\ \hline x_{31} & x_{32} & x_{33} \\ \hline \end{array}, \begin{array}{|c|c|}\hline \frac{\partial p}{\partial y_{11}} & \frac{\partial p}{\partial y_{12}} \\ \hline \frac{\partial p}{\partial y_{21}} & \frac{\partial p}{\partial y_{22}} \\ \hline \end{array} \right)$$

Convolution backward pass

$$\frac{\partial p}{\partial w_{11}} = \frac{\partial p}{\partial y_{11}} x_{11} + \frac{\partial p}{\partial y_{12}} x_{12} + \frac{\partial p}{\partial y_{21}} x_{21} + \frac{\partial p}{\partial y_{22}} x_{22}$$

$$\frac{\partial p}{\partial w_{12}} = \frac{\partial p}{\partial y_{11}} x_{12} + \frac{\partial p}{\partial y_{12}} x_{13} + \frac{\partial p}{\partial y_{21}} x_{22} + \frac{\partial p}{\partial y_{22}} x_{23}$$

$$\frac{\partial p}{\partial w_{21}} = \frac{\partial p}{\partial y_{11}} x_{21} + \frac{\partial p}{\partial y_{12}} x_{22} + \frac{\partial p}{\partial y_{21}} x_{31} + \frac{\partial p}{\partial y_{22}} x_{32}$$

$$\frac{\partial p}{\partial w_{22}} = \frac{\partial p}{\partial y_{11}} x_{22} + \frac{\partial p}{\partial y_{12}} x_{23} + \frac{\partial p}{\partial y_{21}} x_{32} + \frac{\partial p}{\partial y_{22}} x_{33}$$

$$\begin{array}{|c|c|}\hline \frac{\partial p}{\partial w_{11}} & \frac{\partial p}{\partial w_{12}} \\ \hline \frac{\partial p}{\partial w_{21}} & \frac{\partial p}{\partial w_{22}} \\ \hline \end{array} = \mathrm{conv}\left(\begin{array}{|c|c|c|}\hline x_{11} & x_{12} & x_{13} \\ \hline x_{21} & x_{22} & x_{23} \\ \hline x_{31} & x_{32} & x_{33} \\ \hline \end{array}, \begin{array}{|c|c|}\hline \frac{\partial p}{\partial y_{11}} & \frac{\partial p}{\partial y_{12}} \\ \hline \frac{\partial p}{\partial y_{21}} & \frac{\partial p}{\partial y_{22}} \\ \hline \end{array} \right)$$

Convolution backward pass

$$\frac{\partial p}{\partial w_{11}} = \frac{\partial p}{\partial y_{11}} x_{11} + \frac{\partial p}{\partial y_{12}} x_{12} + \frac{\partial p}{\partial y_{21}} x_{21} + \frac{\partial p}{\partial y_{22}} x_{22}$$

$$\frac{\partial p}{\partial w_{12}} = \frac{\partial p}{\partial y_{11}} x_{12} + \frac{\partial p}{\partial y_{12}} x_{13} + \frac{\partial p}{\partial y_{21}} x_{22} + \frac{\partial p}{\partial y_{22}} x_{23}$$

$$\frac{\partial p}{\partial w_{21}} = \frac{\partial p}{\partial y_{11}} x_{21} + \frac{\partial p}{\partial y_{12}} x_{22} + \frac{\partial p}{\partial y_{21}} x_{31} + \frac{\partial p}{\partial y_{22}} x_{32}$$

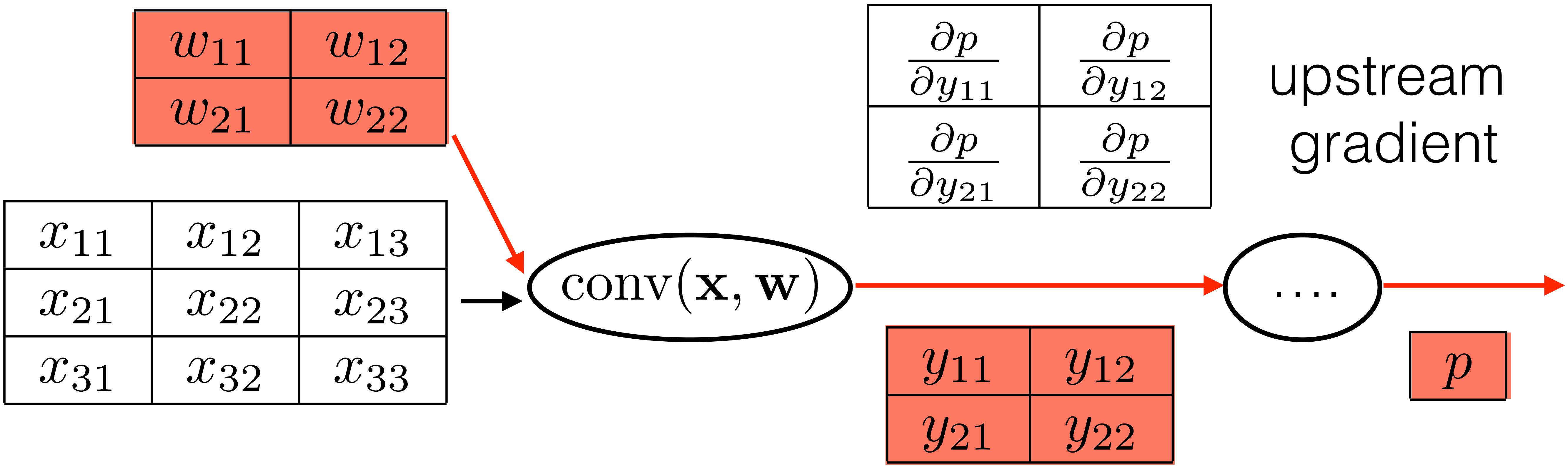
$$\frac{\partial p}{\partial w_{22}} = \frac{\partial p}{\partial y_{11}} x_{22} + \frac{\partial p}{\partial y_{12}} x_{23} + \frac{\partial p}{\partial y_{21}} x_{32} + \frac{\partial p}{\partial y_{22}} x_{33}$$

$$\begin{array}{|c|c|}\hline \frac{\partial p}{\partial w_{11}} & \frac{\partial p}{\partial w_{12}} \\ \hline \frac{\partial p}{\partial w_{21}} & \frac{\partial p}{\partial w_{22}} \\ \hline \end{array} = \mathrm{conv}\left(\begin{array}{|c|c|c|}\hline x_{11} & x_{12} & x_{13} \\ \hline x_{21} & x_{22} & x_{23} \\ \hline x_{31} & x_{32} & x_{33} \\ \hline \end{array}, \begin{array}{|c|c|}\hline \frac{\partial p}{\partial y_{11}} & \frac{\partial p}{\partial y_{12}} \\ \hline \frac{\partial p}{\partial y_{21}} & \frac{\partial p}{\partial y_{22}} \\ \hline \end{array} \right)$$

Convolution backward pass wrt weights

- Backpropagation in convolutional layer wrt weights is: "**convolution of input feature map with upstream gradient**"

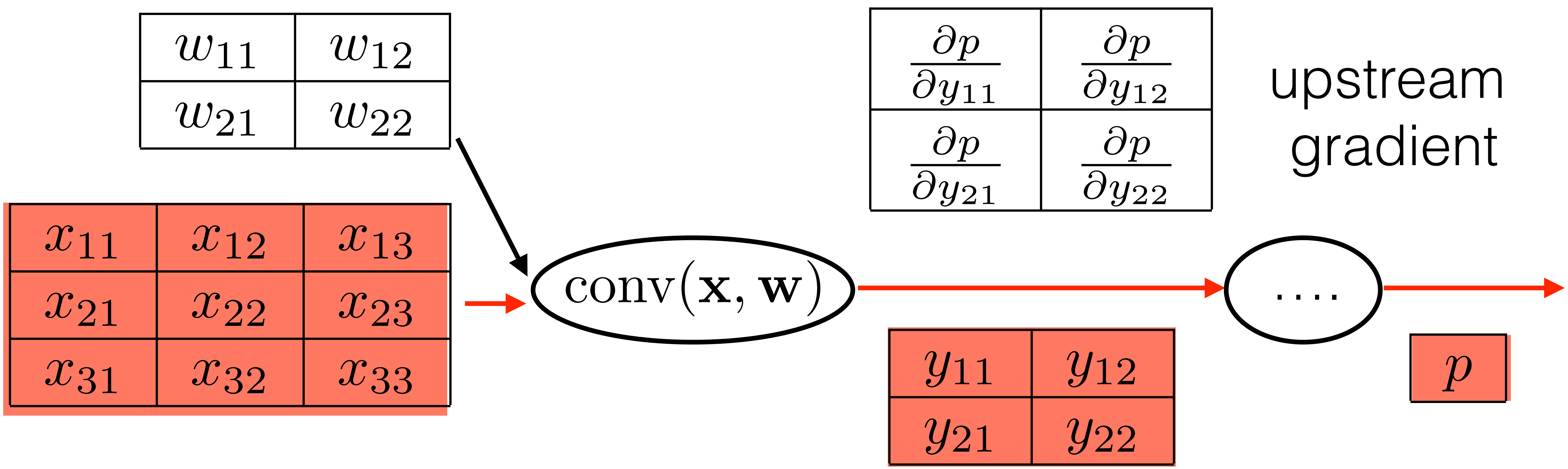
$$\begin{bmatrix} \frac{\partial p}{\partial w_{11}} & \frac{\partial p}{\partial w_{12}} \\ \frac{\partial p}{\partial w_{21}} & \frac{\partial p}{\partial w_{22}} \end{bmatrix} = \mathrm{conv}\left(\begin{bmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \end{bmatrix}, \begin{bmatrix} \frac{\partial p}{\partial y_{11}} & \frac{\partial p}{\partial y_{12}} \\ \frac{\partial p}{\partial y_{21}} & \frac{\partial p}{\partial y_{22}} \end{bmatrix} \right)$$



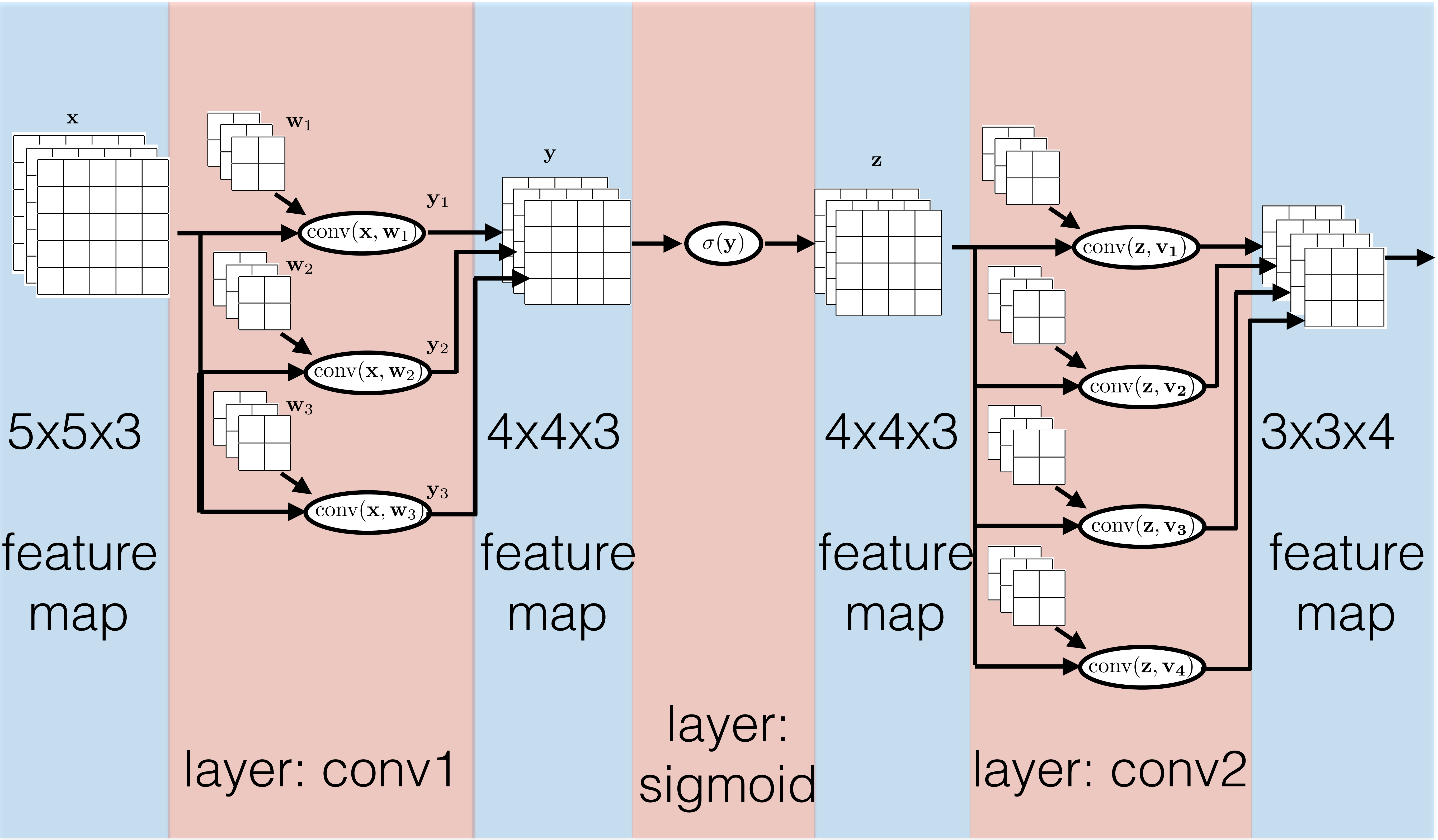
Convolution backward pass wrt input feature map

Backpropagation in convolutional layer is:
"convolution of padded upstream gradient with mirrored weights"

$$\begin{bmatrix} \frac{\partial p}{\partial x_{11}} & \frac{\partial p}{\partial x_{12}} & \frac{\partial p}{\partial x_{13}} \\ \frac{\partial p}{\partial x_{21}} & \frac{\partial p}{\partial x_{22}} & \frac{\partial p}{\partial x_{23}} \\ \frac{\partial p}{\partial x_{31}} & \frac{\partial p}{\partial x_{32}} & \frac{\partial p}{\partial x_{33}} \end{bmatrix} = \mathrm{conv}\left(\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & \frac{\partial p}{\partial y_{11}} & \frac{\partial p}{\partial y_{12}} & 0 \\ 0 & \frac{\partial p}{\partial y_{21}} & \frac{\partial p}{\partial y_{22}} & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} w_{22} & w_{21} \\ w_{12} & w_{11} \end{bmatrix} \right)$$



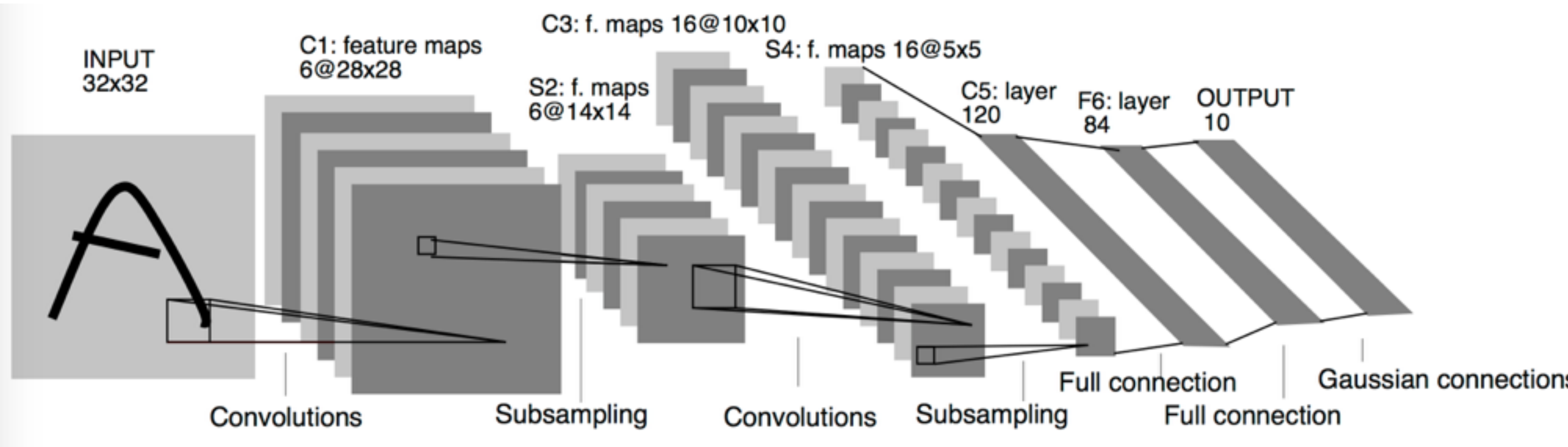
???? Convolutional network backprop ?????



Convolutional net

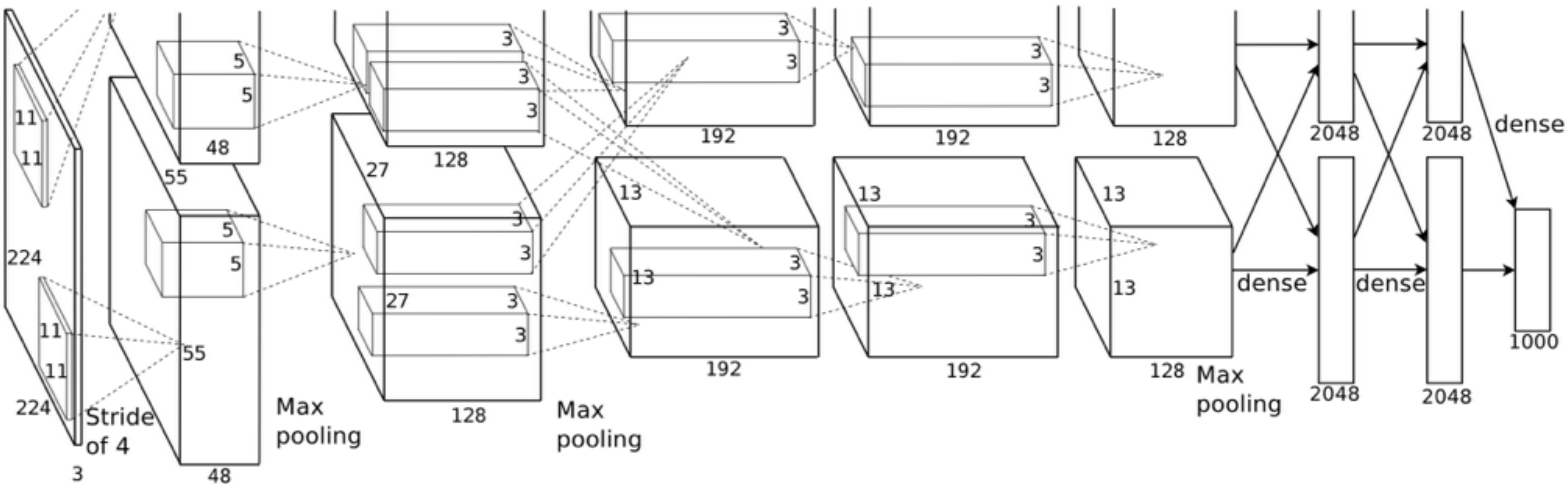
- Convolutional network (ConvNet) is concatenation of convolutional layers
- Backprop in ConvNet is convolution of feature maps or kernels with upstream gradient.
- Feed-forward and backprop are concolutions => efficient implementation on GPU

LeCun's letter recognition 1998 (over 13k citations !!!)



LeCun et al, Gradient based learning applied to document recognition, IEEE, 1998
http://yann.lecun.com/exdb/publis/pdf/lecun-01a.pdf

AlexNet on ImageNet 2012 (**over 27k citations !!!**)



Alex Krizhevsky et al, Imagenet classification with deep convolutional neural networks, NIPS, 2012
https://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networks.pdf

IMAGENET

Classification results

http://image-net.org/challenges/LSVRC/2017/index

Steel drum

Output:
Scale
T-shirt
<u>Steel drum</u>
Drumstick
Mud turtle

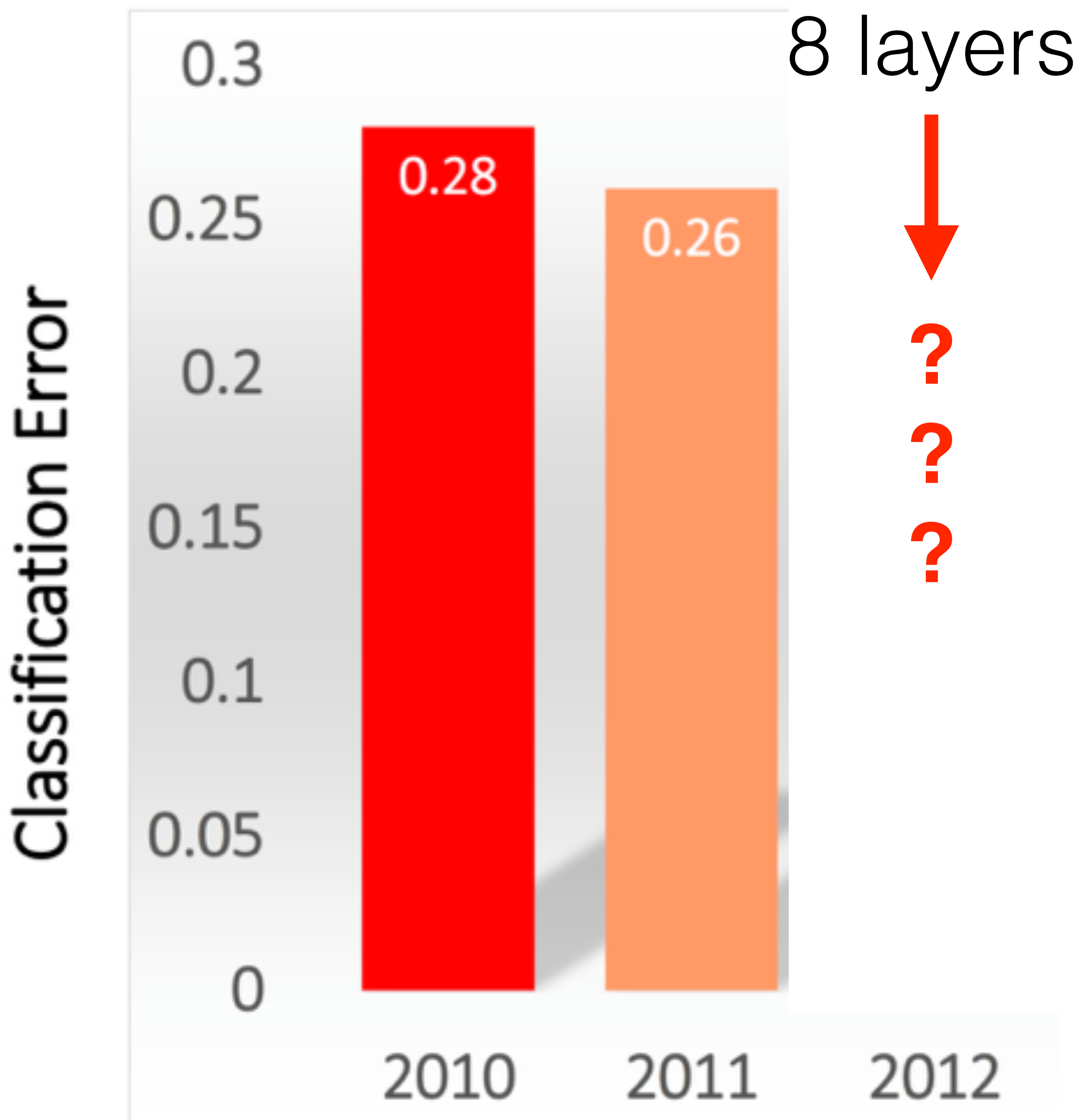
Output:
Scale
T-shirt
Giant panda
Drumstick
Mud turtle

$$\text{Error} = \frac{1}{100{,}000} \sum_{\substack{100{,}000 \\ \text{images}}} 1[\text{incorrect on image } i]$$

IMAGENET

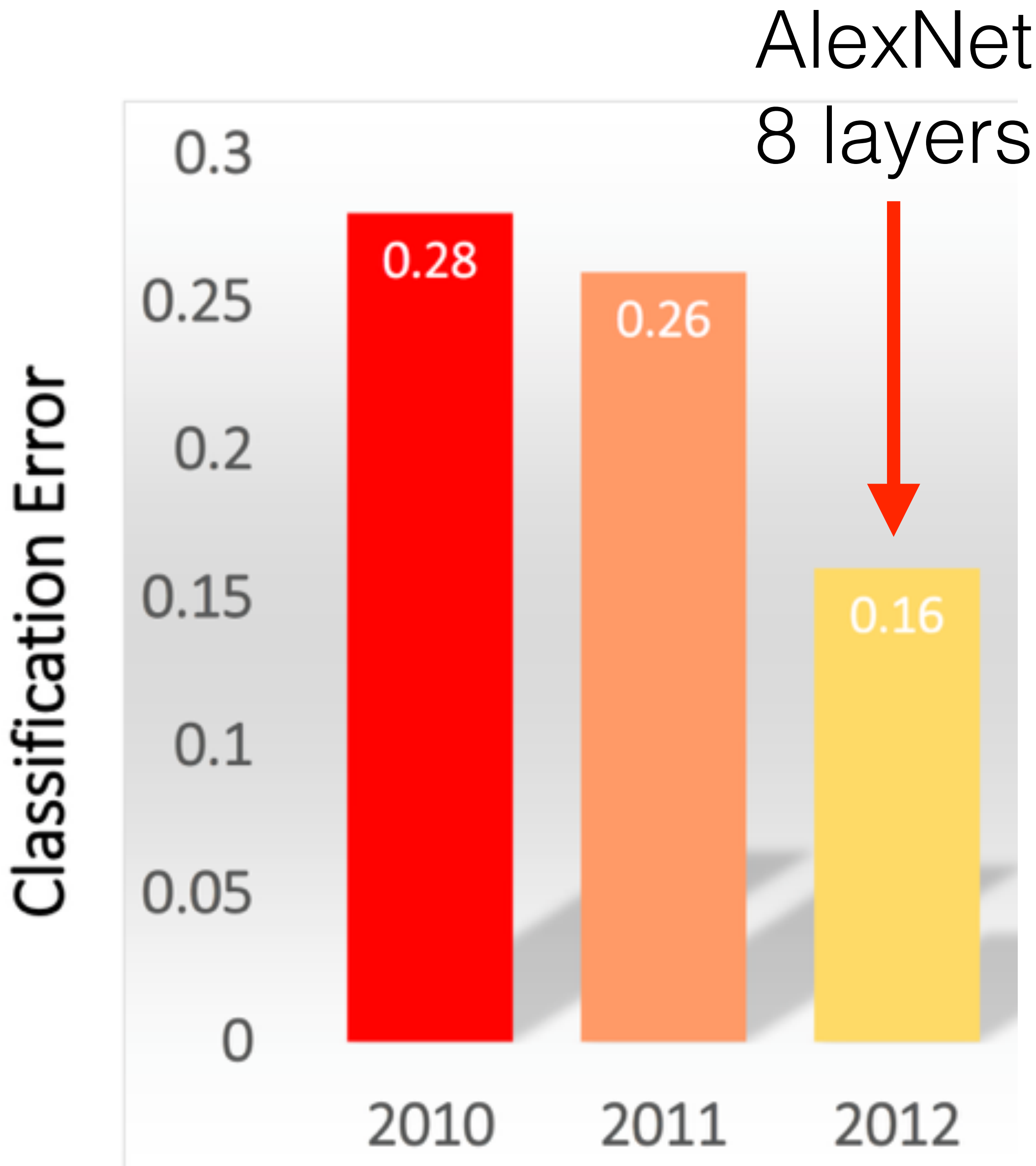
Classification results

AlexNet
8 layers



IMAGENET

Classification results

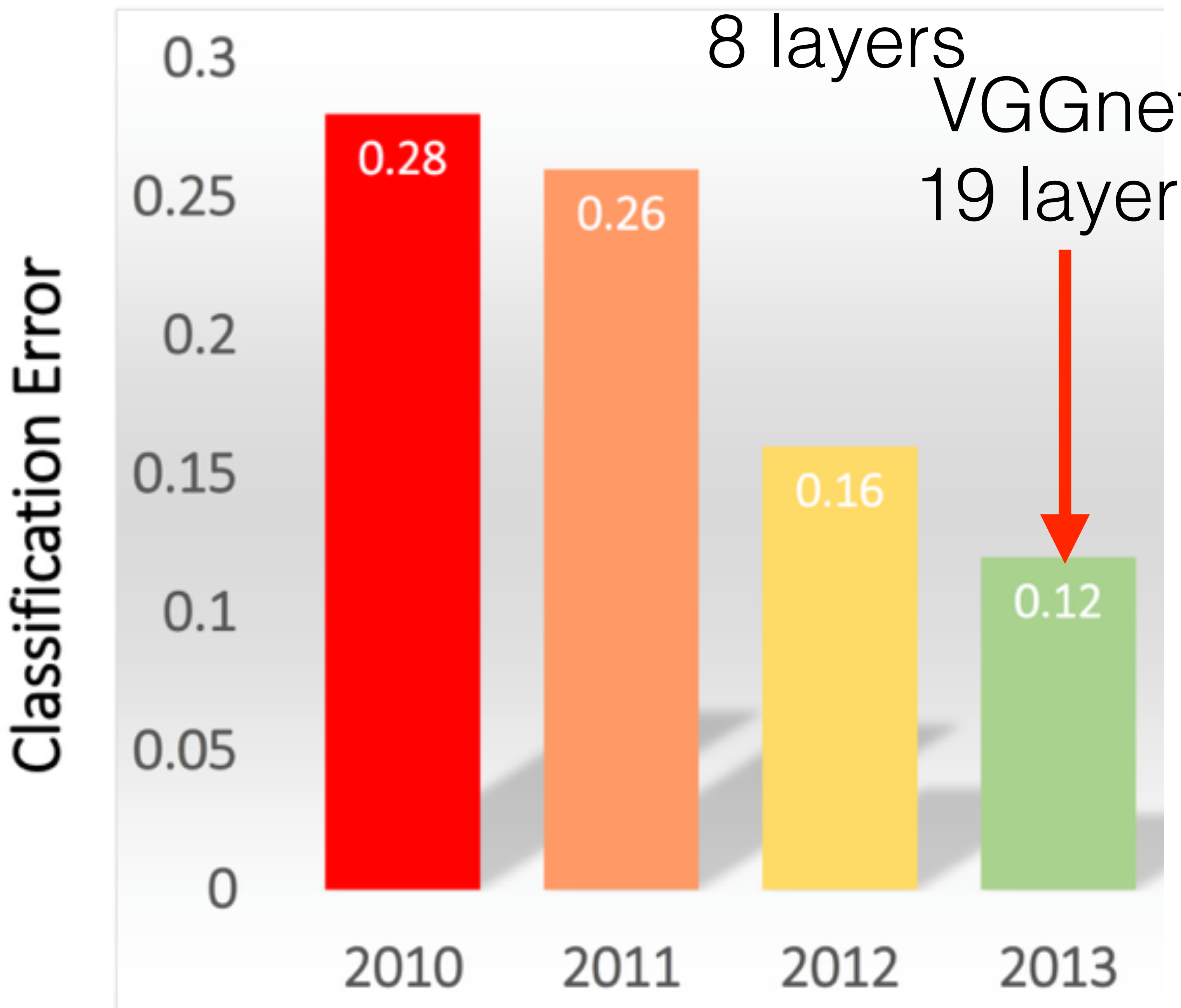


IMAGENET

Classification results

AlexNet
8 layers

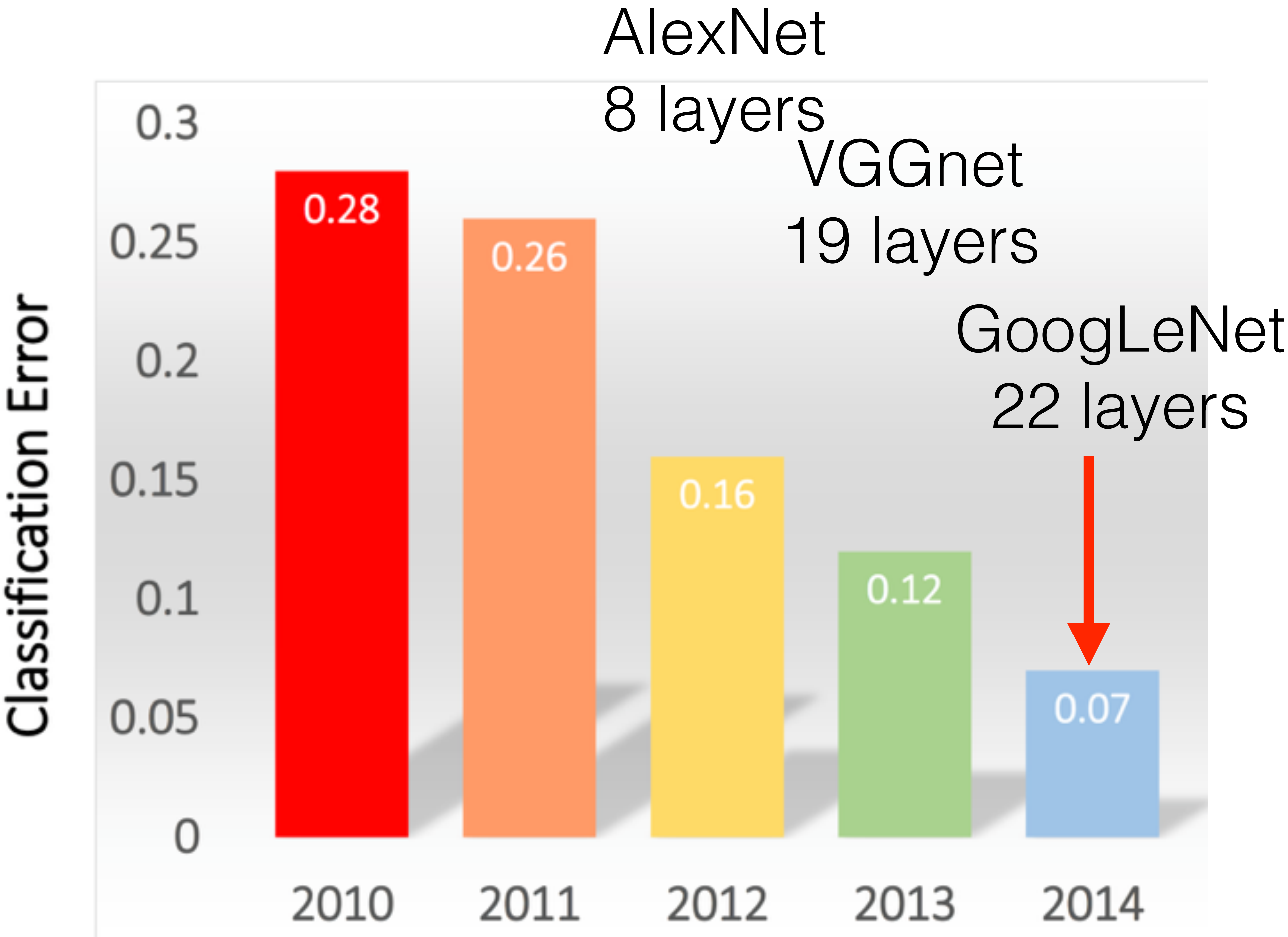
VGGnet
19 layers



Czech Technical University in Prague
Faculty of Electrical Engineering, Department of Cybernetics

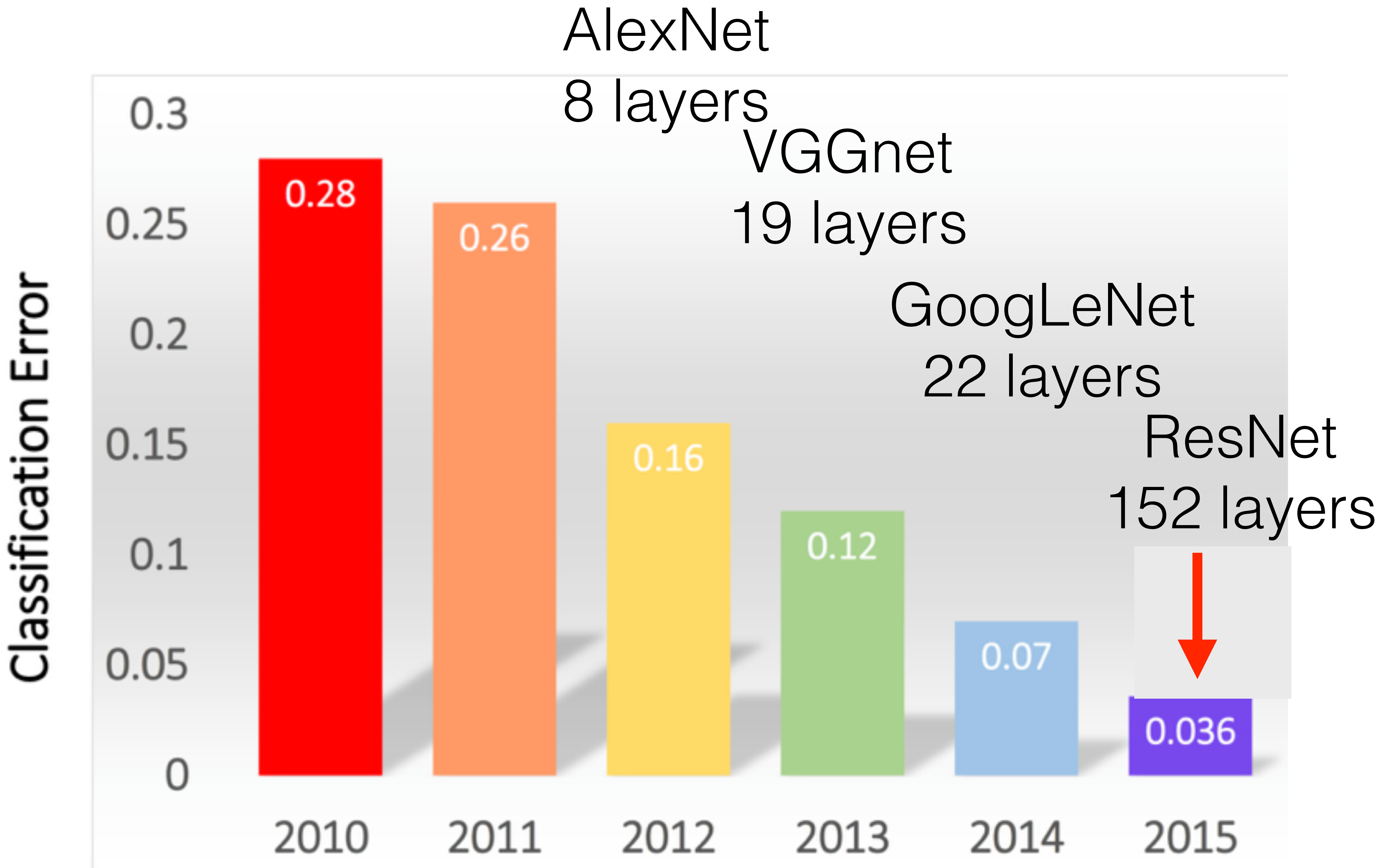
IMAGENET

Classification results



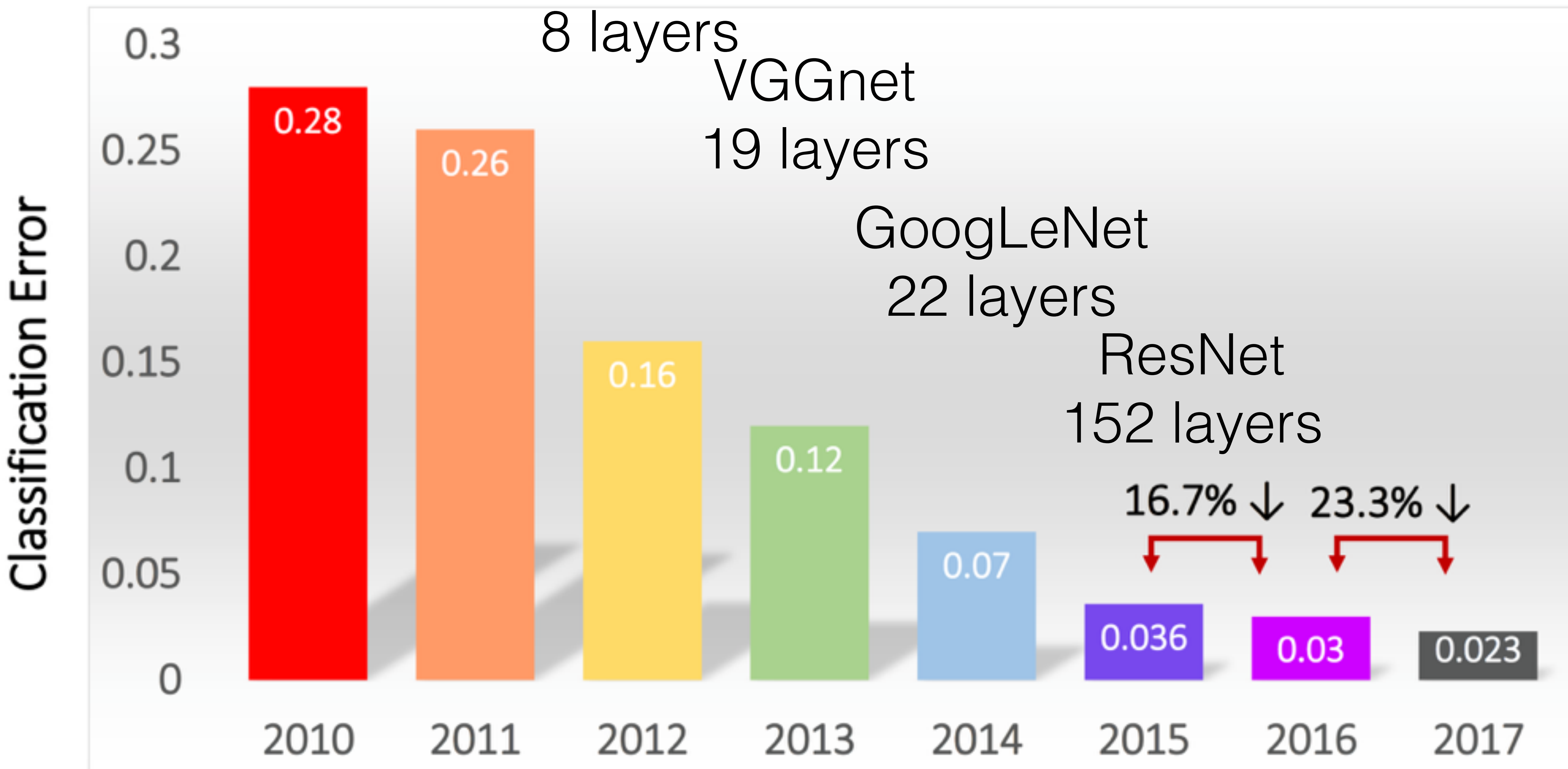
IMAGENET

Classification results



IMAGENET

Classification results



Demo

- convnet demo from Karpathy: https://cs.stanford.edu/people/karpathy/convnetjs/demo/cifar10.html

Next lecture

- gradient learning (what make it tough)
- other layers:
  - activation function,
  - batch normalization,
  - drop out,
  - loss layers