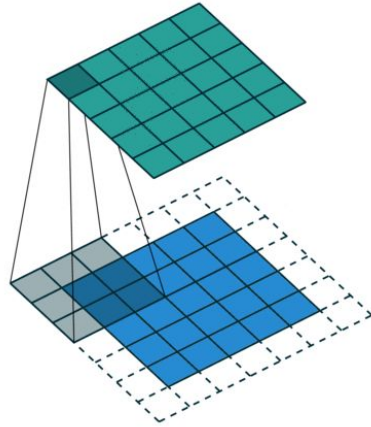
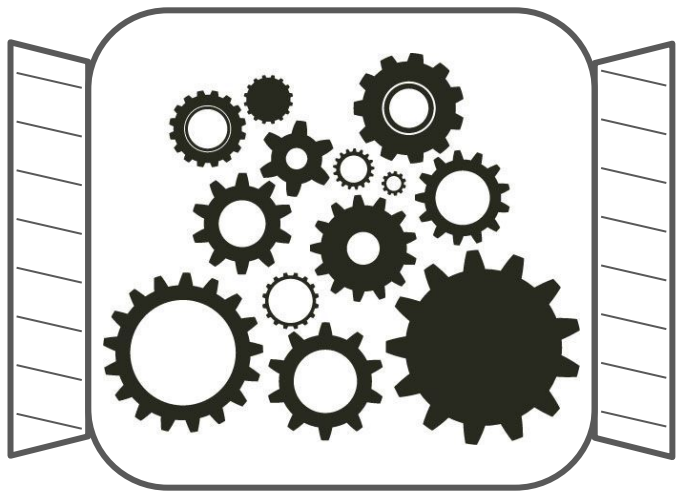


Convolutional Networks

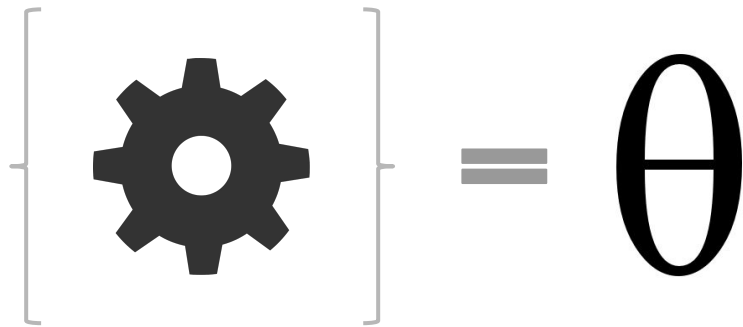


BY BOTOS CSABA

Parametric function



$$f : X \rightarrow Y$$

A diagram illustrating a parametric function. It features a single gear icon enclosed within a large left-facing curly bracket. To the right of the bracket is an equals sign, followed by the Greek letter θ . This visualizes the concept of a function defined by a single set of parameters.

$$f(x) = f_{\theta}(x)$$

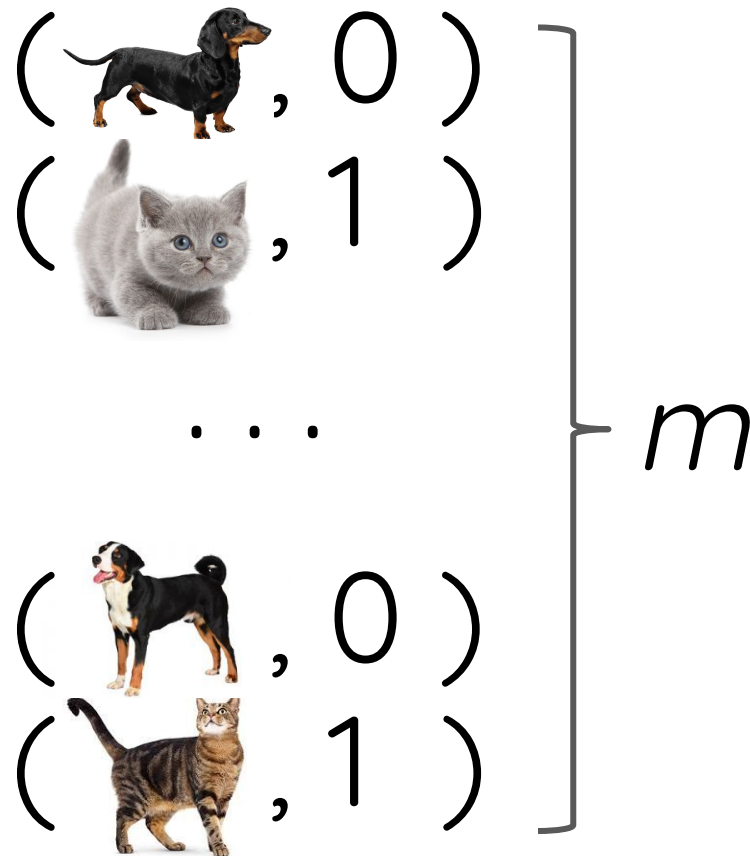
Loss function

Given m number of $(\mathbf{x}, y')$
sample pairs



Loss function

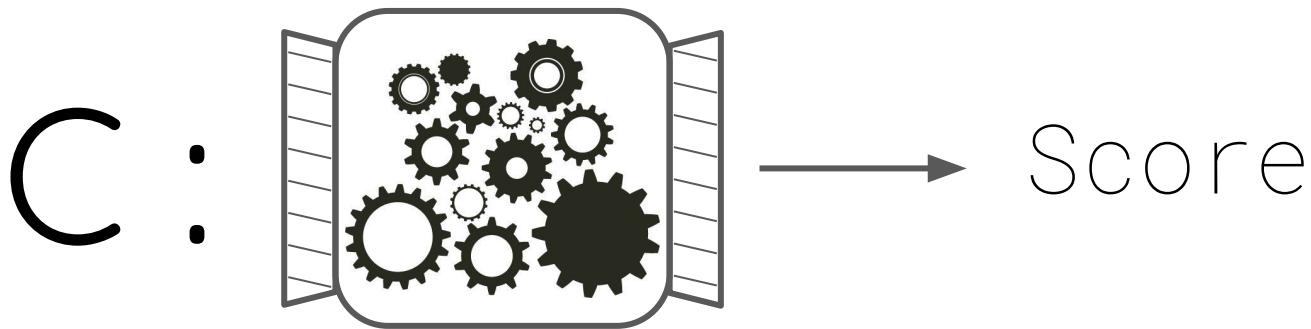
Given m number of $(\mathbf{x}, \mathbf{y}')$
sample pairs



Loss function

Given m number of $(\mathbf{x}, \mathbf{y}')$
sample pairs

The Cost function scores a
function that predicts $\mathbf{y}$

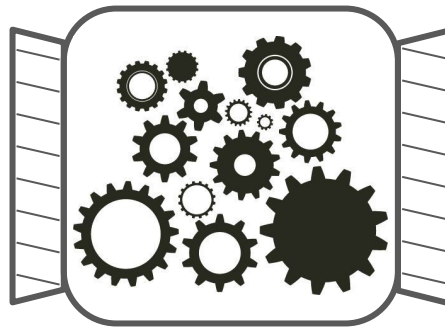


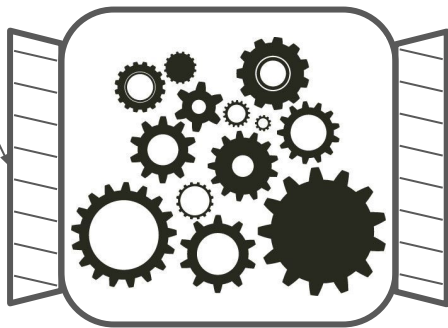


...

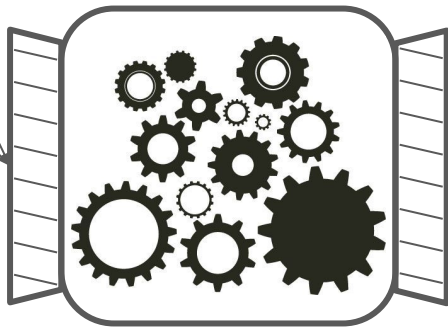


How does it work?





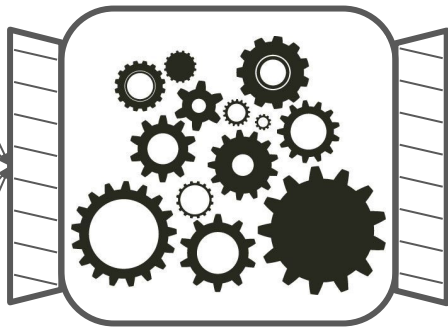
$(0.1, 0)$



$$\begin{pmatrix} 0.1, & 0 \end{pmatrix} \\ \begin{pmatrix} 0.8, & 1 \end{pmatrix}$$



...



$$\begin{pmatrix} 0.1, & 0 \end{pmatrix}$$

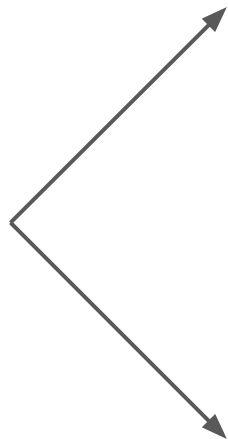
$$\begin{pmatrix} 0.8, & 1 \end{pmatrix}$$

...

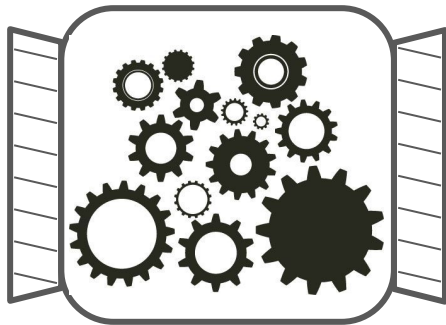
$$\begin{pmatrix} 0.4, & 0 \end{pmatrix}$$

$$\begin{pmatrix} 0.3, & 1 \end{pmatrix}$$

x



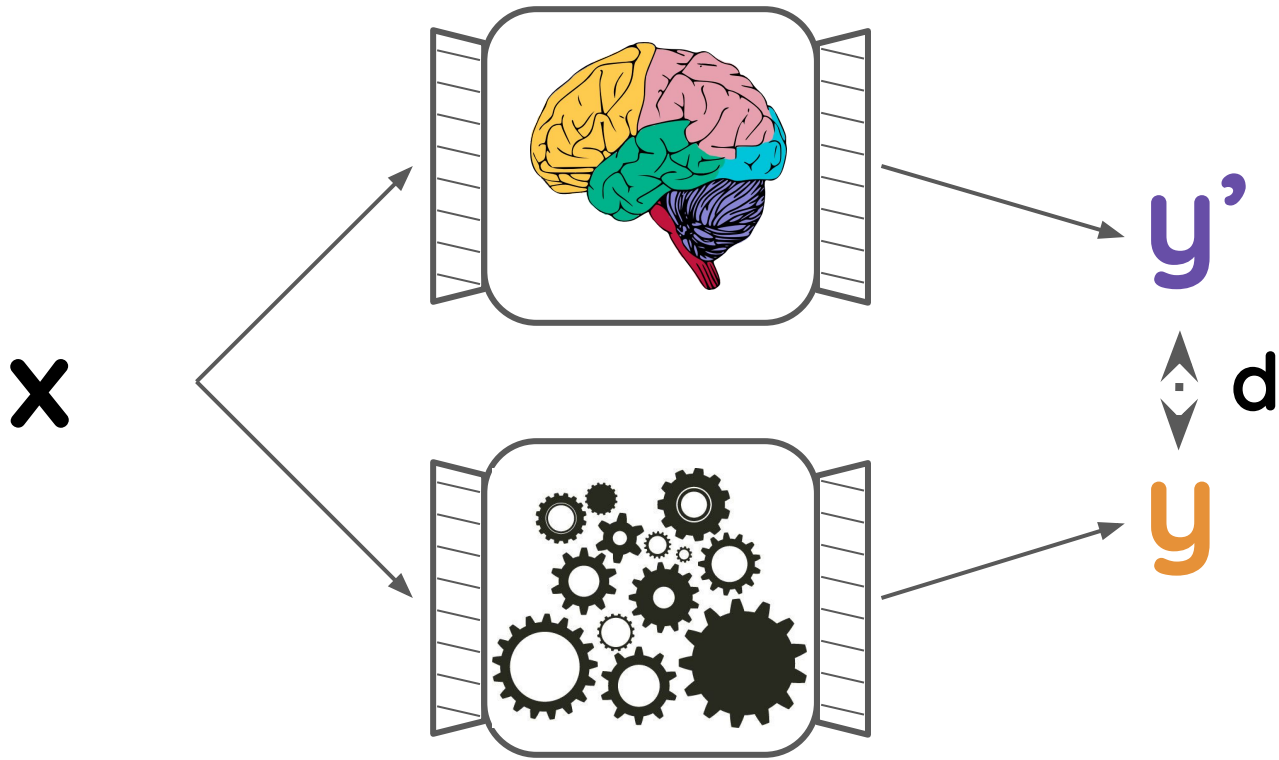
y'

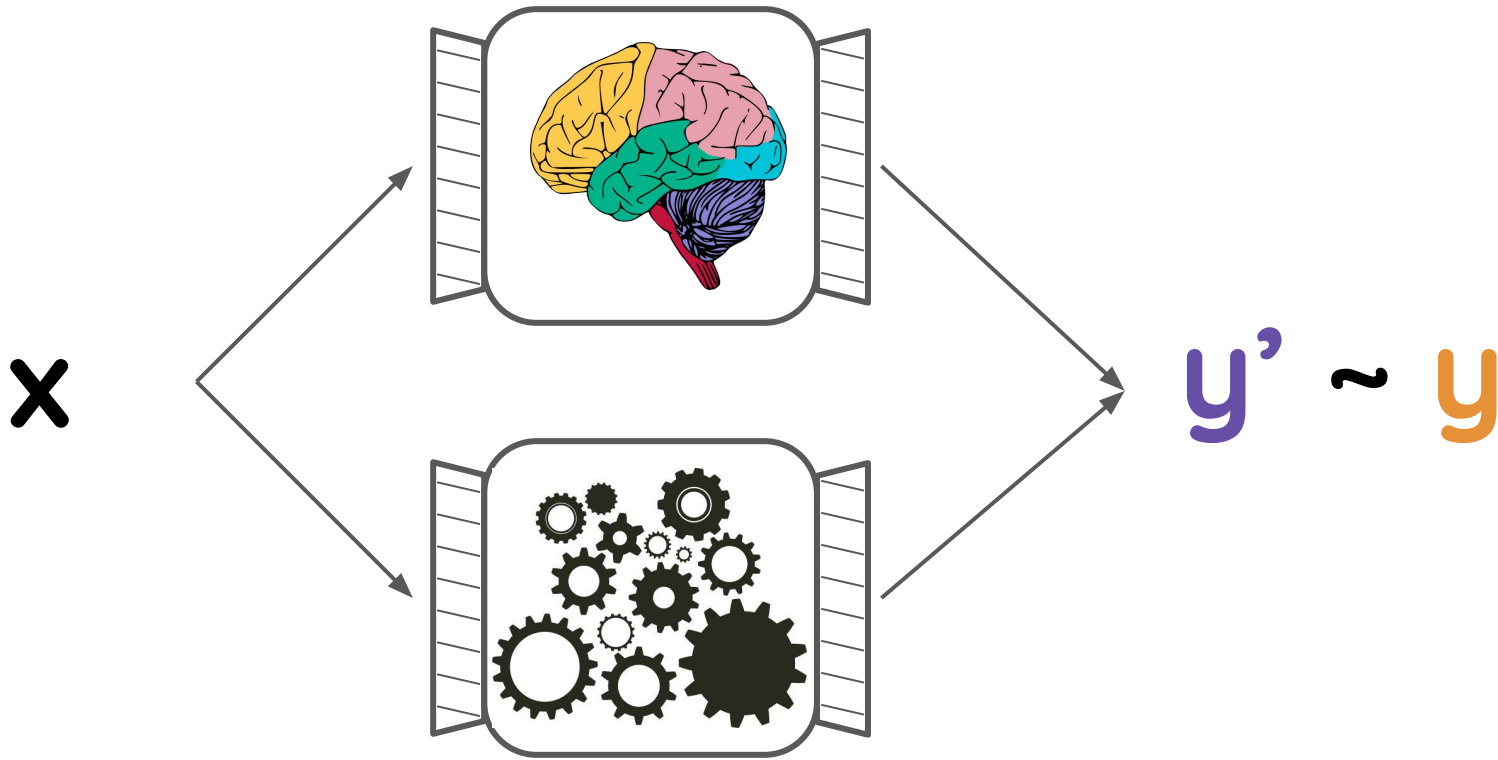


y

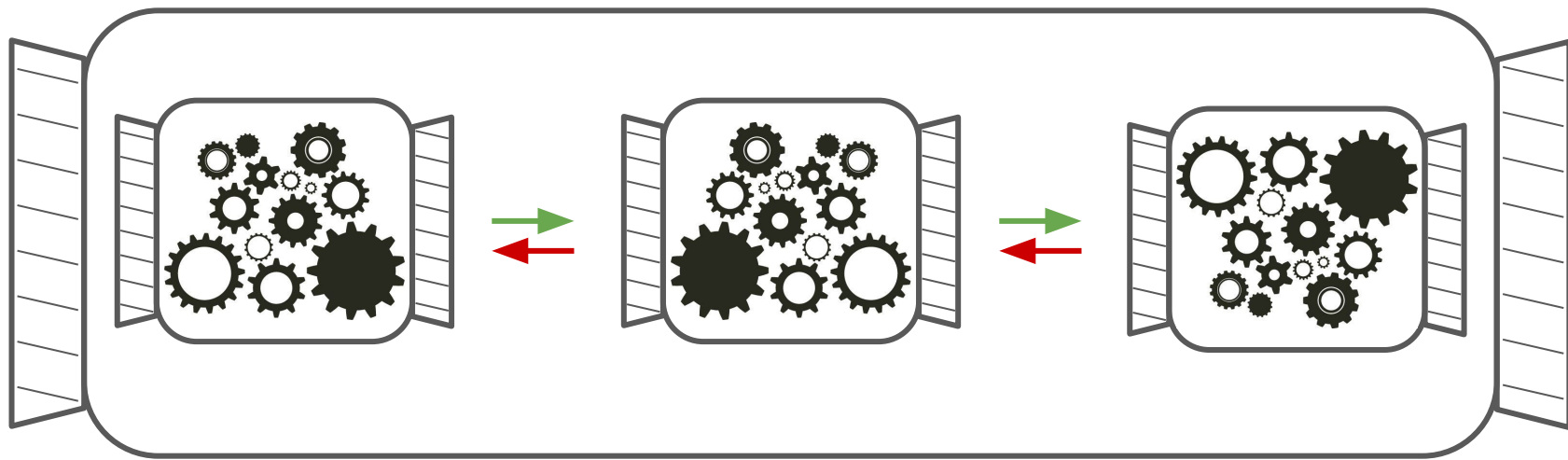


a





Embedded parametric functions



1. Forward

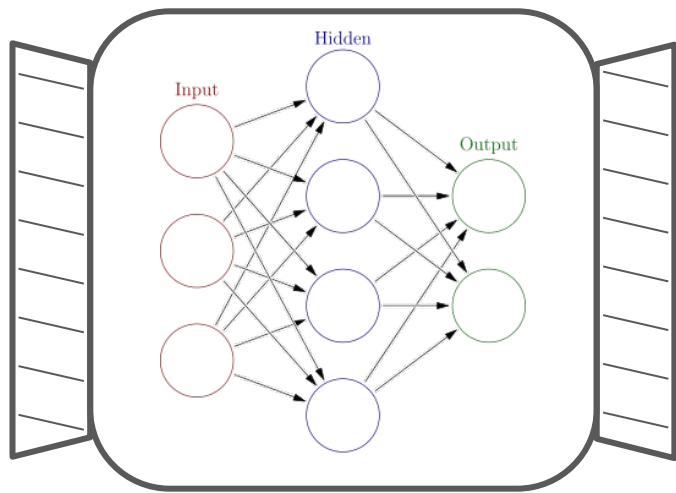


2. Backward



3. Update 

Neural Networks

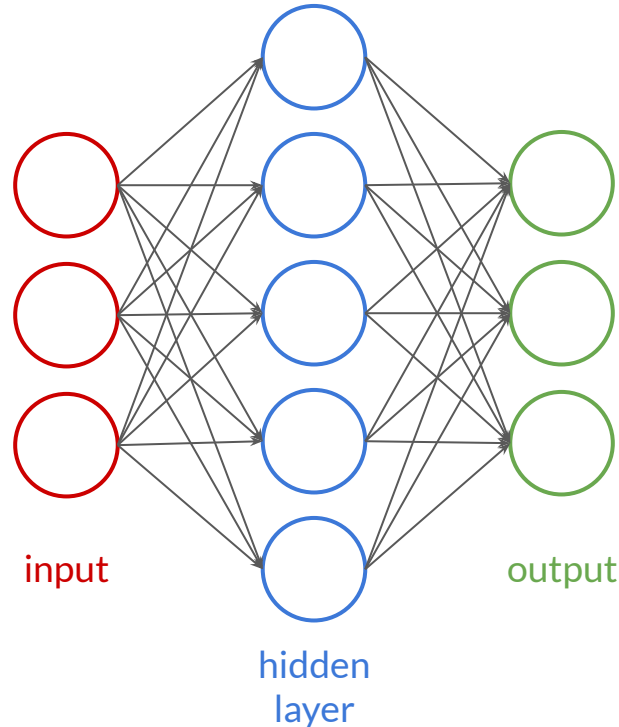


$$f : \mathbf{X} \rightarrow \mathbf{Y}$$

$$\{\mathbf{w}, \mathbf{b}\} = \boldsymbol{\theta}$$

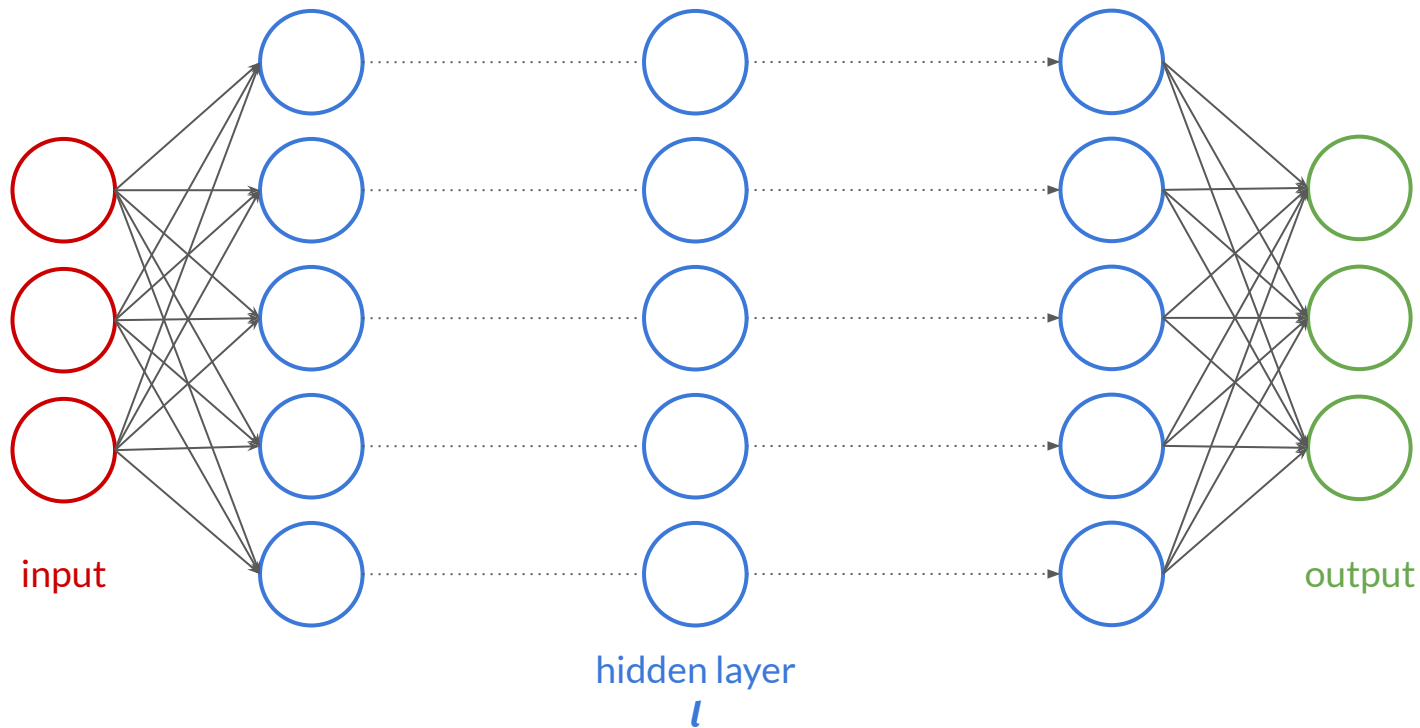
$$f(x) = f_{\theta}(x)$$

2 layer network - everything is possible



$$a_2 (a_1 (x))$$

New layer's definition



1. Forward



2. Backward

3. Update

Image Processing - grayscale

Dimensions:

$[H, W]$



Single convolution

Weight:
[3, 3]

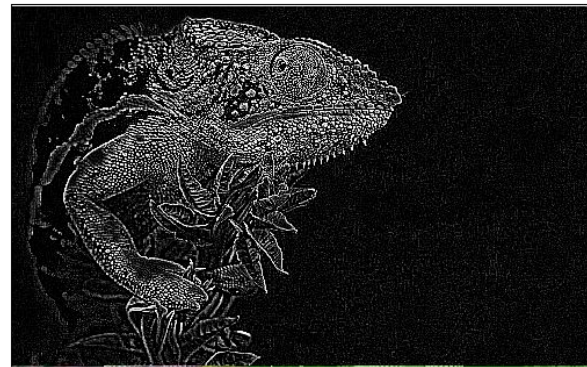


Input:
[H+2, W+2]

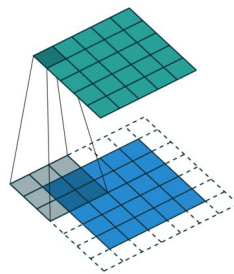
*

-1	-1	-1
-1	8	-1
-1	-1	-1

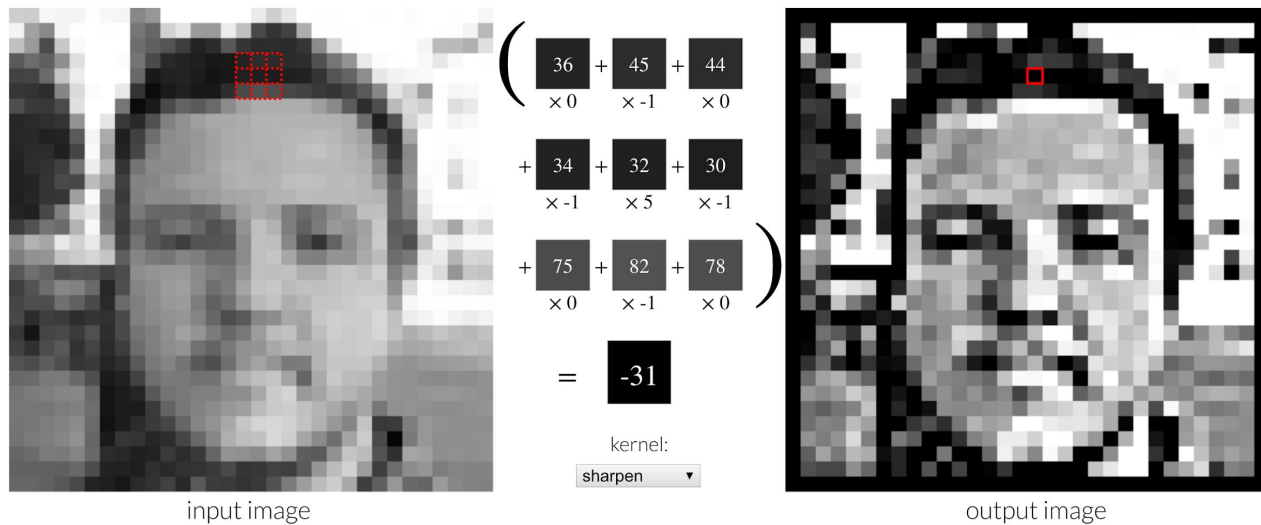
=



Output:
[H, W]



Convolution demo



Live demo: <http://setosa.io/ev/image-kernels/>

Multiple convolution



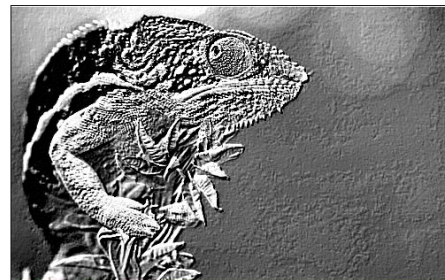
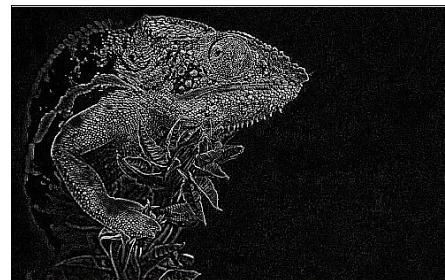
Input:
[H+2, W+2]

*

-1	-1	-1
-1	8	-1
-1	-1	-1

-2	-1	0
-1	1	1
0	1	2

=



Weight:
[3, 3, **2**]

Output:
[H, W, **2**]

Multiple convolution



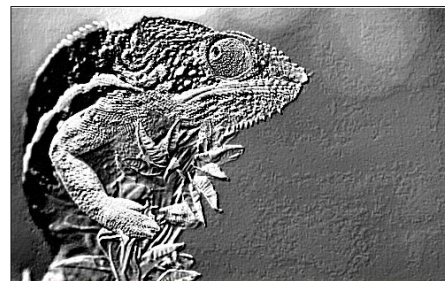
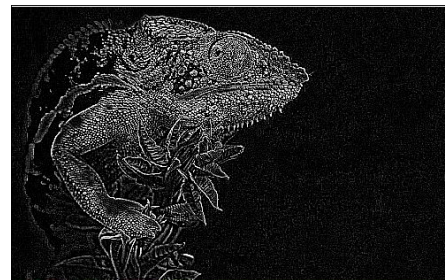
Input:
 $[H+2, W+2]$

*

-1	-1	-1
-1	8	-1
-1	-1	-1

-2	-1	0
-1	1	1
0	1	2

=



Weight:
 $[3, 3, \mathbf{C}_{out}]$

Output:
 $[H, W, \mathbf{C}_{out}]$

Image Processing - multi-channel

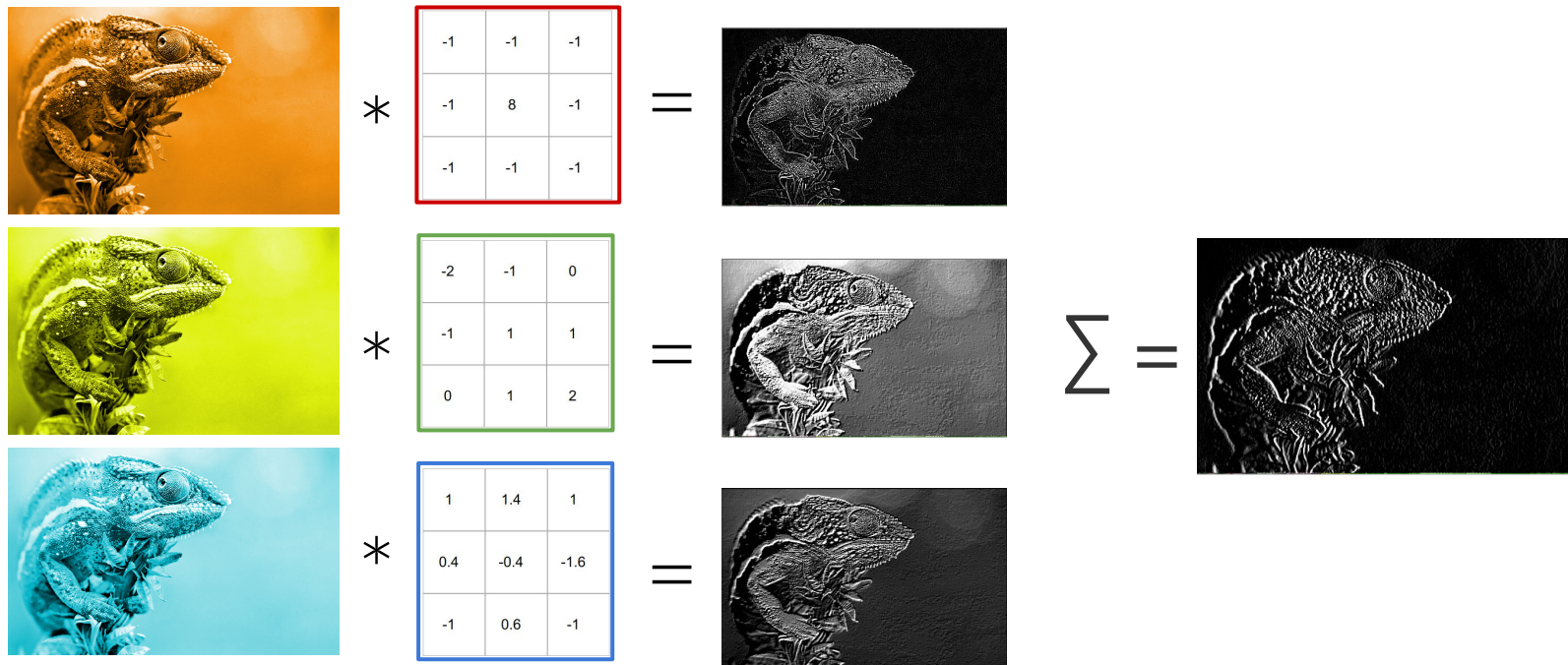
Dimensions:

[H, W, C]

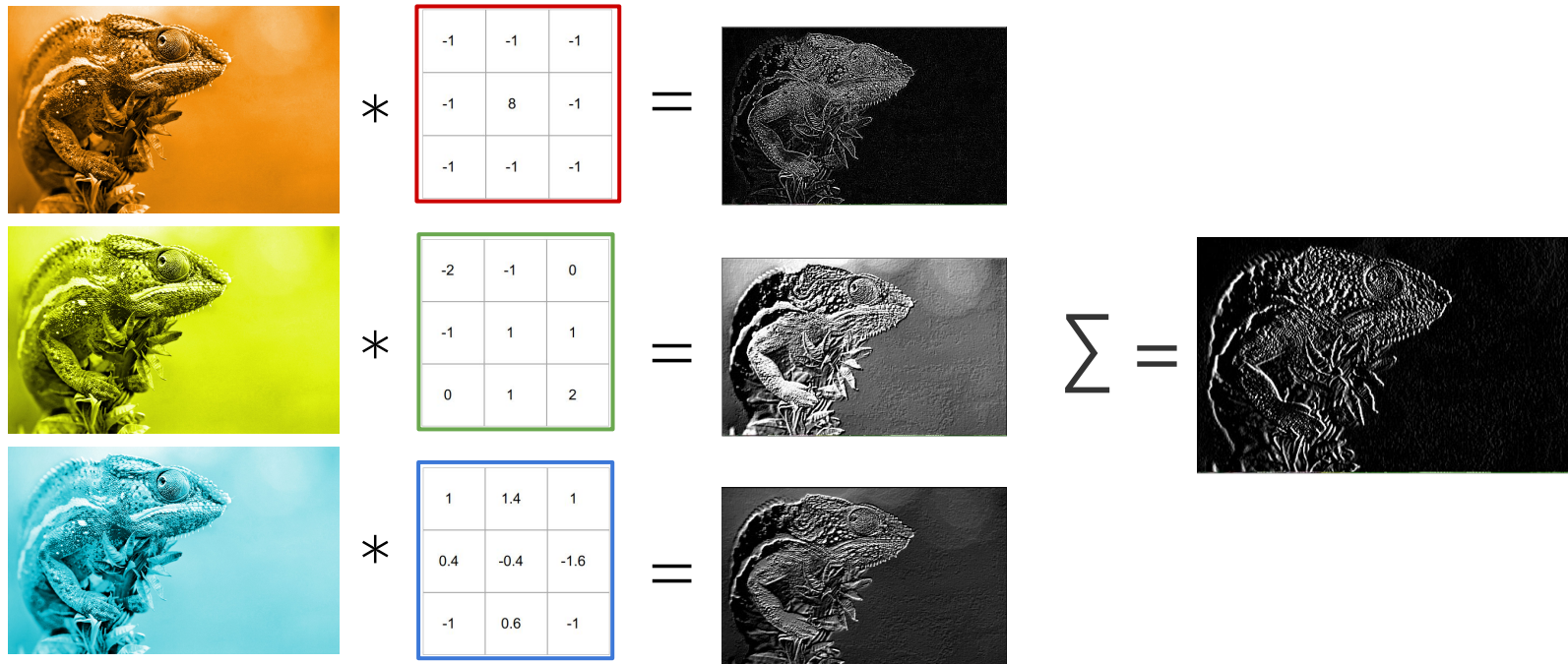


$k = 1..C$

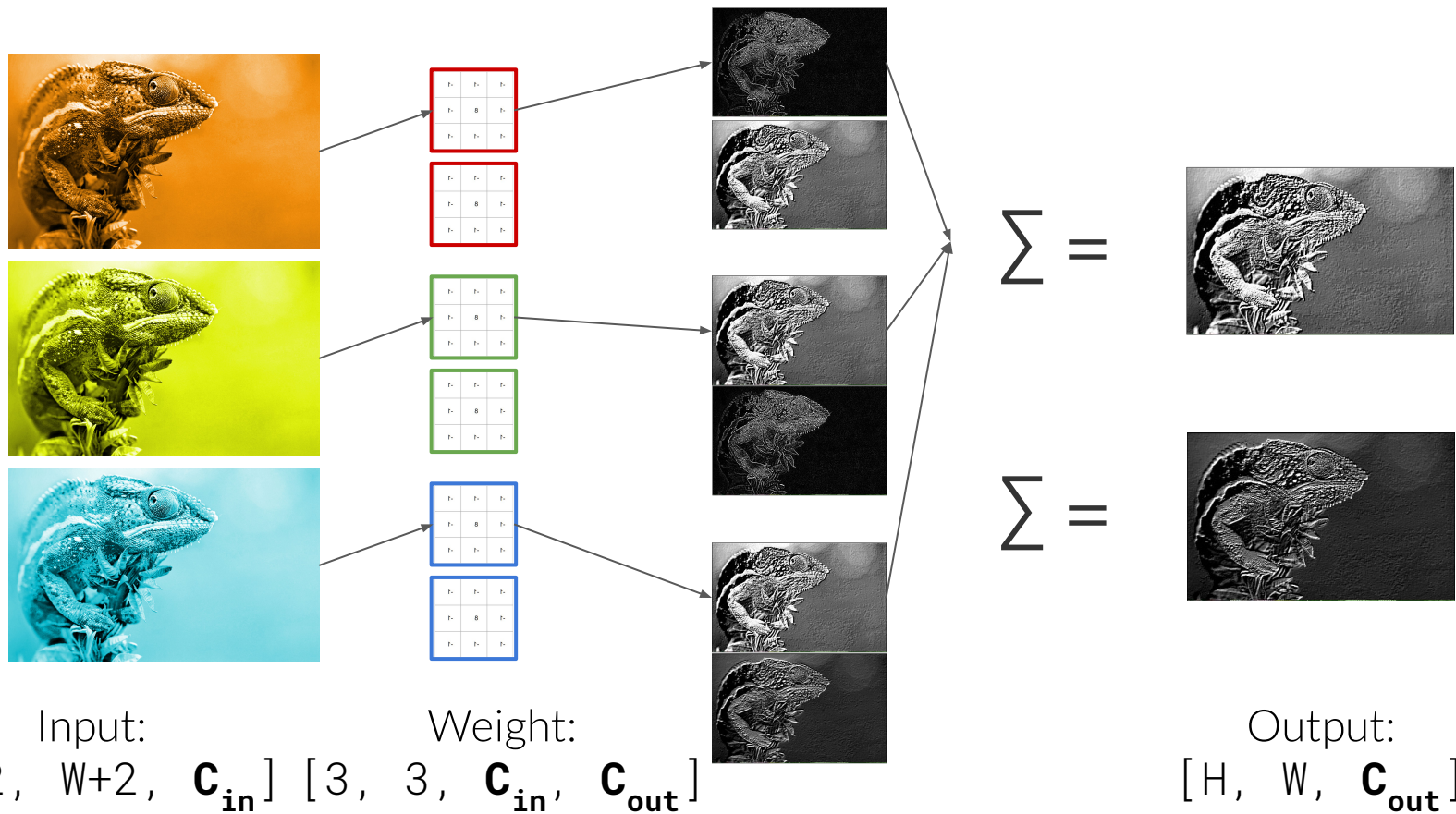
Multi-channel convolution



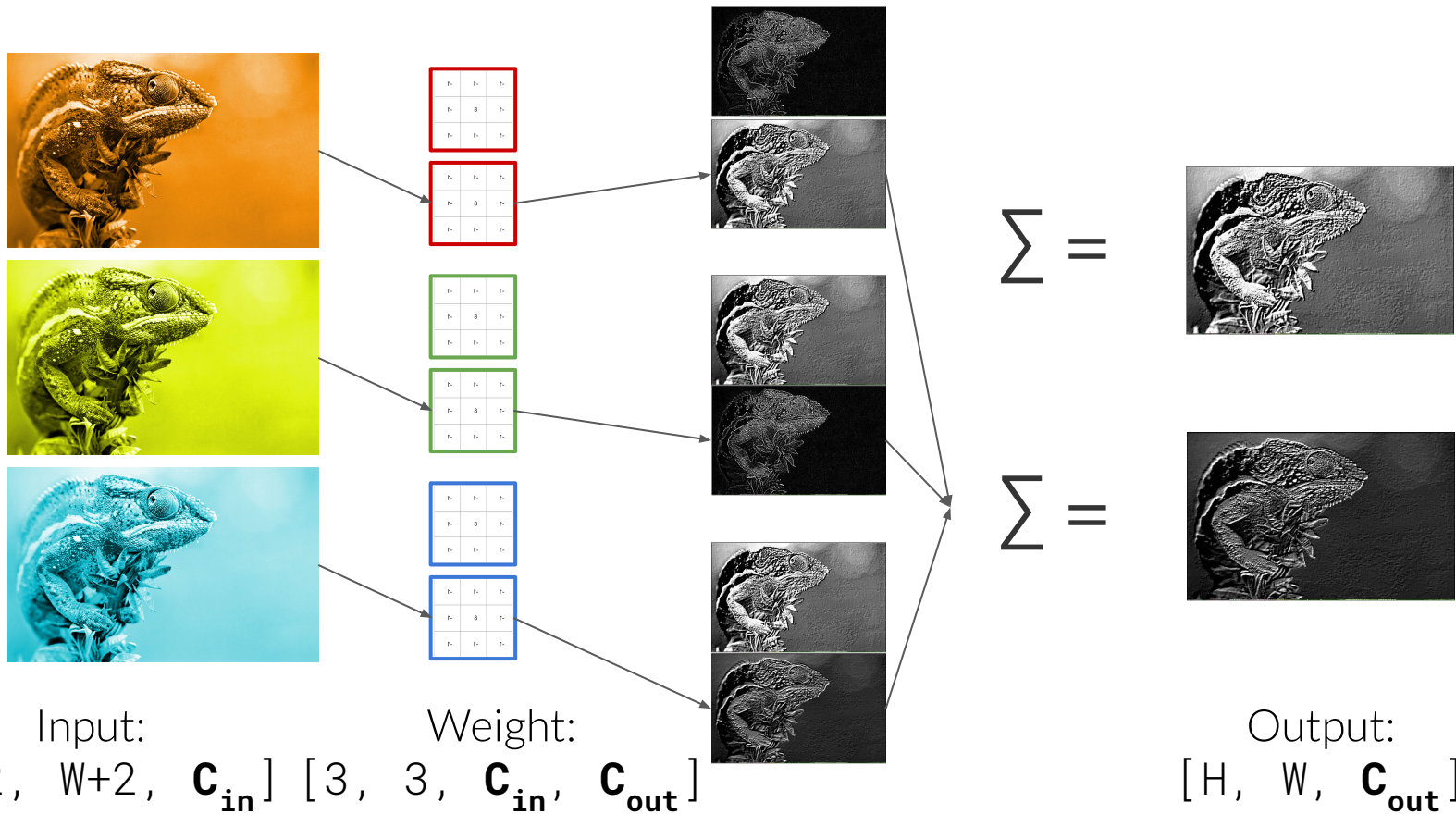
Multi-channel convolution



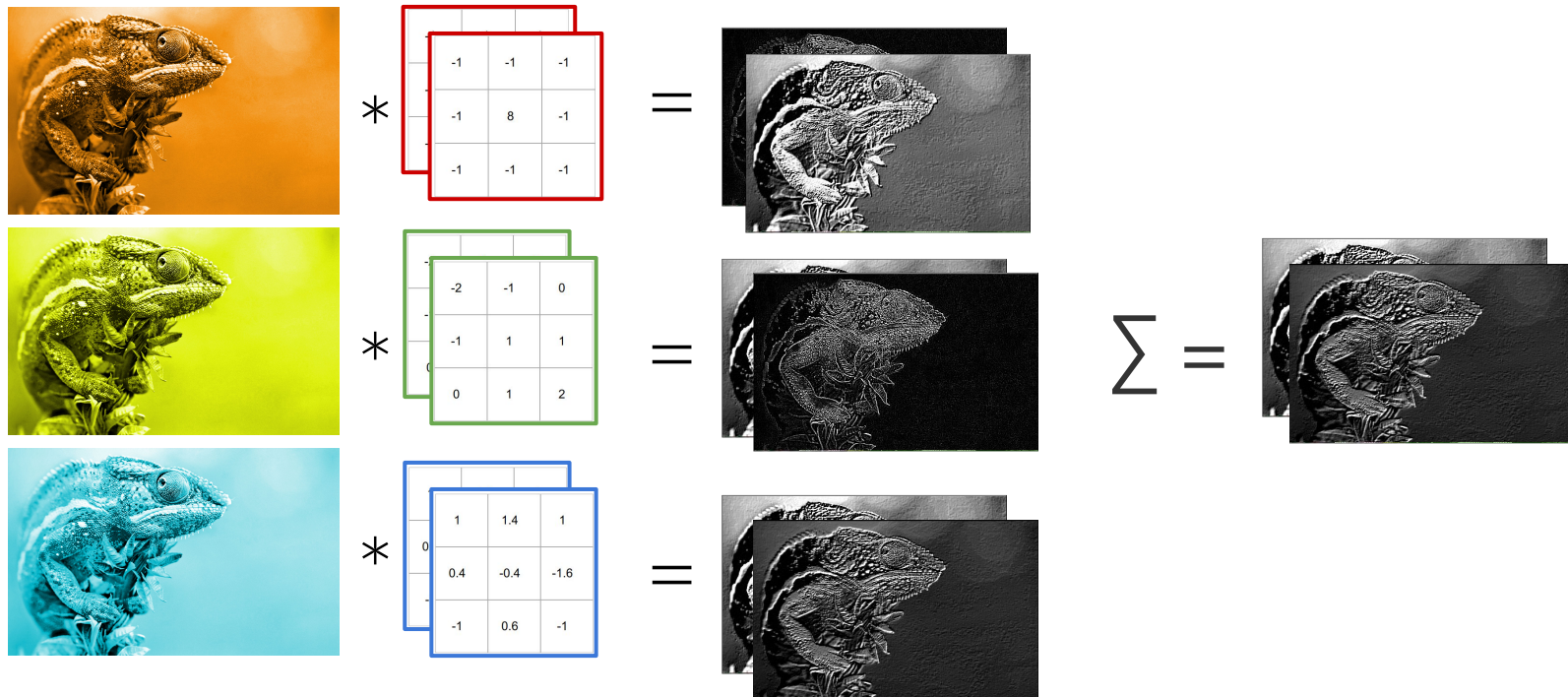
General convolution



General convolution



General convolution

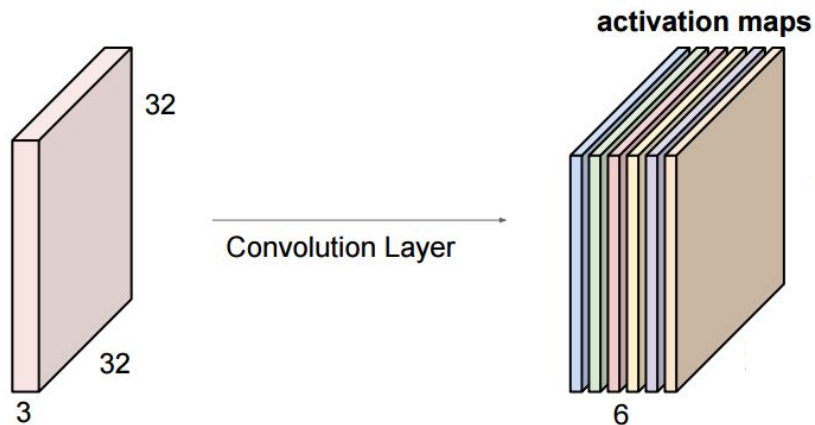
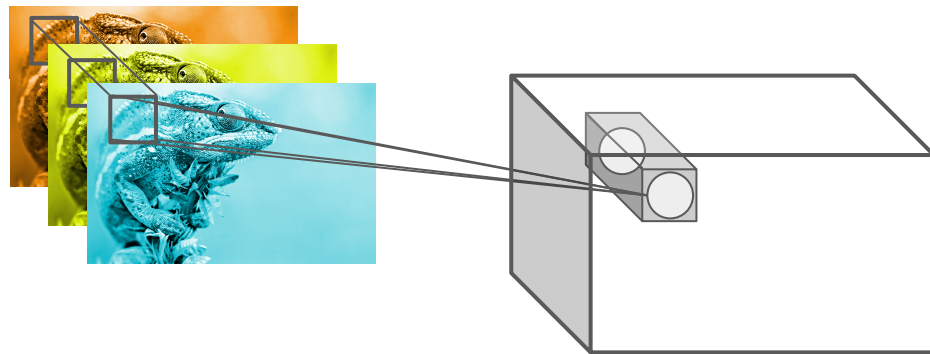
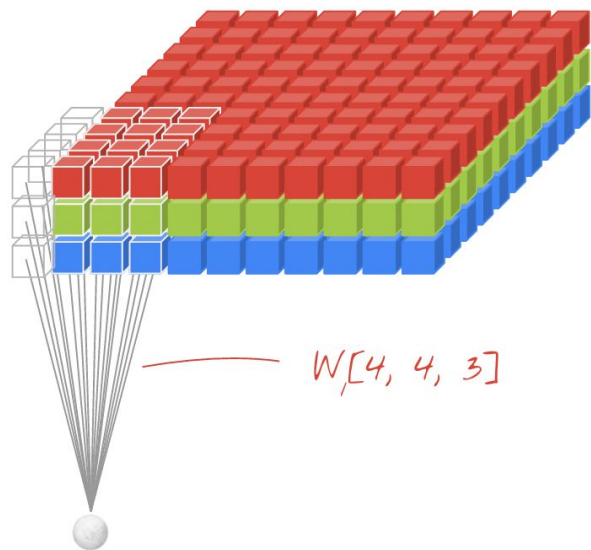


Input:
 $[H+2, W+2, \mathbf{C}_{in}]$

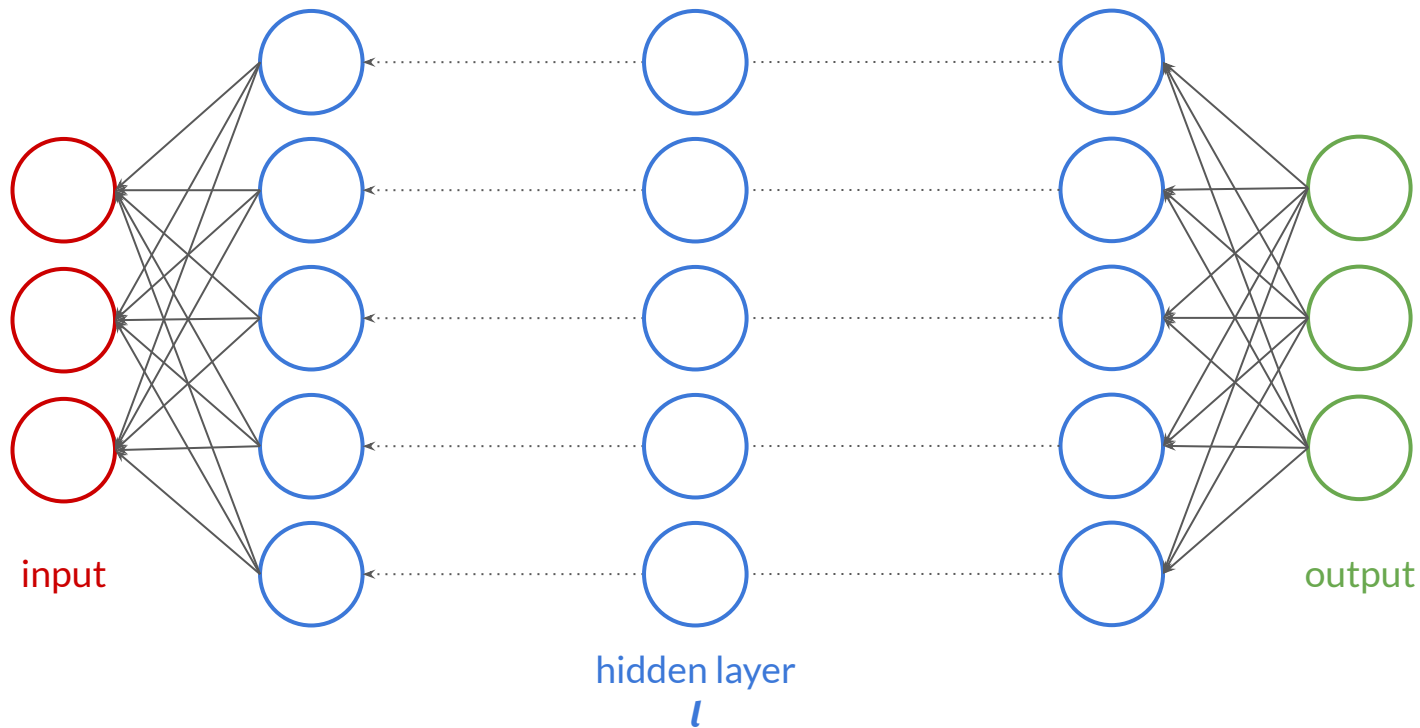
Weight:
 $[3, 3, \mathbf{C}_{in}, \mathbf{C}_{out}]$

Output:
 $[H, W, \mathbf{C}_{out}]$

Convolutional Layer



New layer's definition



1. Forward



2. Backward



3. Update

Single convolution - backward

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

$$\delta^{l-1} = \frac{\partial y}{\partial x} \delta^l$$

Single convolution - backward

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

$$\frac{\partial y}{\partial x_{2,2}} = ?$$

In this case:

x_{22} has influence on every output

Single convolution - backward

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

w_{33}

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

Single convolution - backward

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

w_{33} w_{32}

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

Single convolution - backward

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

w_{33} w_{32} w_{31}

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

Single convolution - backward

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

w_{33} w_{32} w_{31}

w_{23}

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

Single convolution - backward

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

w_{33} w_{32} w_{31}
 w_{23} w_{22}

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

Single convolution - backward

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

w_{33} w_{32} w_{31}

w_{23} w_{22} w_{21}

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

Single convolution - backward

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

w_{33} w_{32} w_{31}

w_{23} w_{22} w_{21}

w_{13}

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

Single convolution - backward

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

w_{33} w_{32} w_{31}

w_{23} w_{22} w_{21}

w_{13} w_{12}

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

Single convolution - backward

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

w_{33} w_{32} w_{31}

w_{23} w_{22} w_{21}

w_{13} w_{12} w_{11}

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

Single convolution - backward

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

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

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

Single convolution - backward

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

forward kernel
= $\mathbf{W}$

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

backward kernel
= $\mathbf{W}[:, :, -1, ::-1]$

Single convolution - backward

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

$$\delta^{l-1} = \frac{\partial y}{\partial x} \delta^l$$

Single convolution - backward

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

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

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

Single convolution - backward

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

$$\delta^{l-1} =$$

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

$$\frac{\partial y}{\partial x} * \delta^l$$

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

Single convolution - backward

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

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

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

Single convolution - backward

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

$$\delta^{l-1} =$$

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

$$\frac{\partial y}{\partial x} * \delta^l$$

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

Single convolution - FORWARD

forward kernel
= **W**

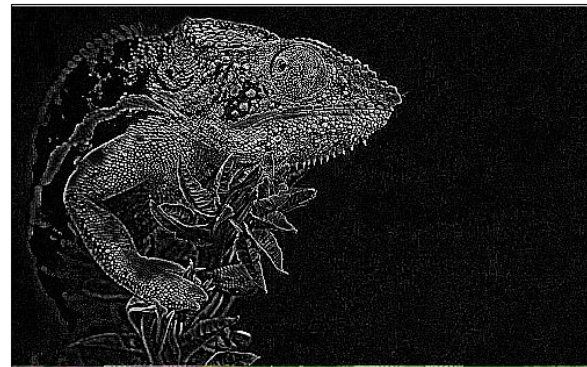


δ^{l-1}
[H+2, W+2]

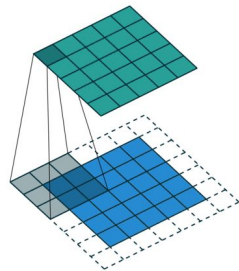
*

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

=

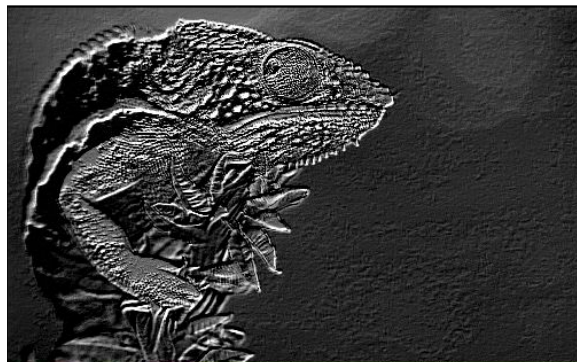


δ^l
[H, W]



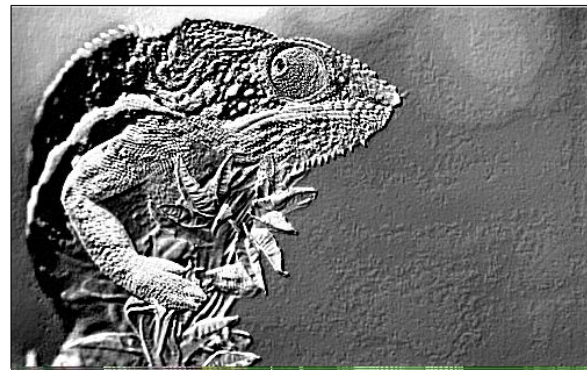
Single convolution - BACKWARD

backward kernel
= $\mathbf{W}[:, :, -1, : -1]$

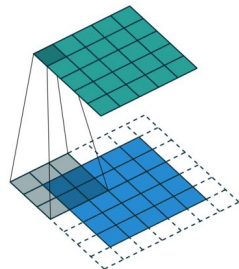


δ^{l-1}
[H+2, W+2]

$$= \begin{bmatrix} w_{33} & w_{32} & w_{31} \\ w_{23} & w_{22} & w_{21} \\ w_{13} & w_{12} & w_{11} \end{bmatrix} *$$



δ^l
[H, W]



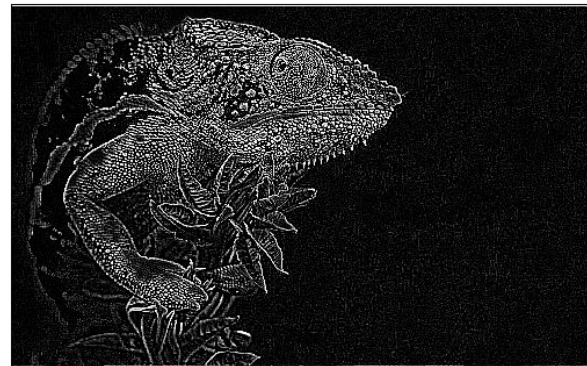
Forward



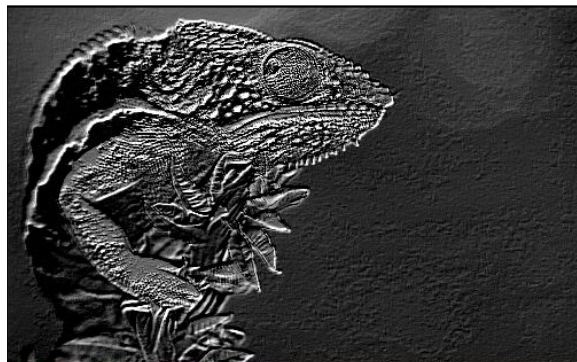
*

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

=



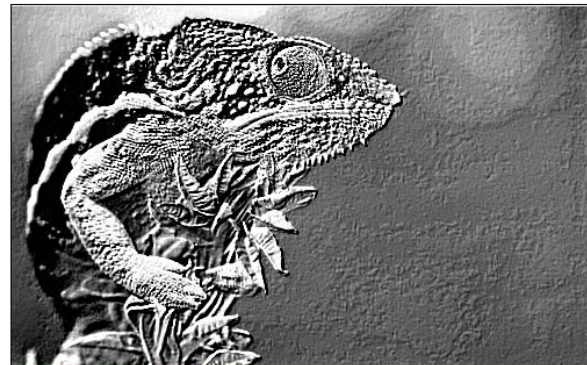
Backward



=

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

*



Forward



$$X * W = Y$$

Backward



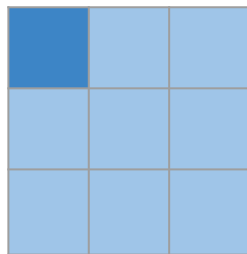
$$\delta^{l-1} = W^{180^\circ} * \delta^l$$

Multiple convolution - FORWARD

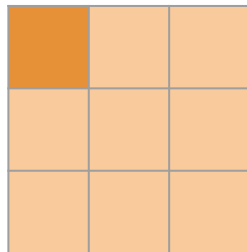


Input:
 $[H+2, W+2]$

*



=



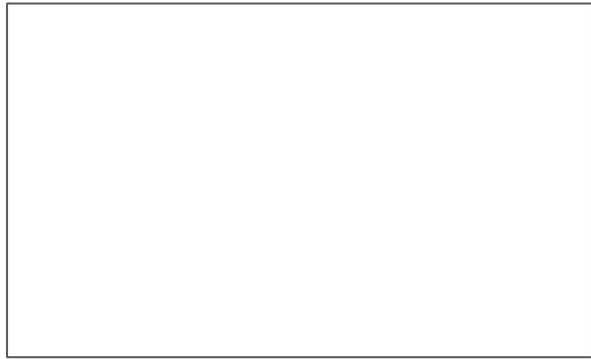
W
 $[3, 3, \mathbf{C}_{out}]$



Output:
 $[H, W, \mathbf{C}_{out}]$

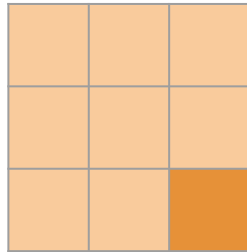
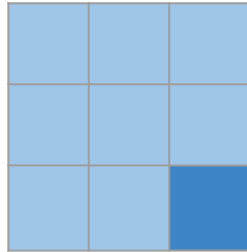
Multiple convolution - BACKWARD

Accumulate errors



$$\delta^{l-1}$$
$$[H+2, W+2, \mathbf{1}]$$

= Σ



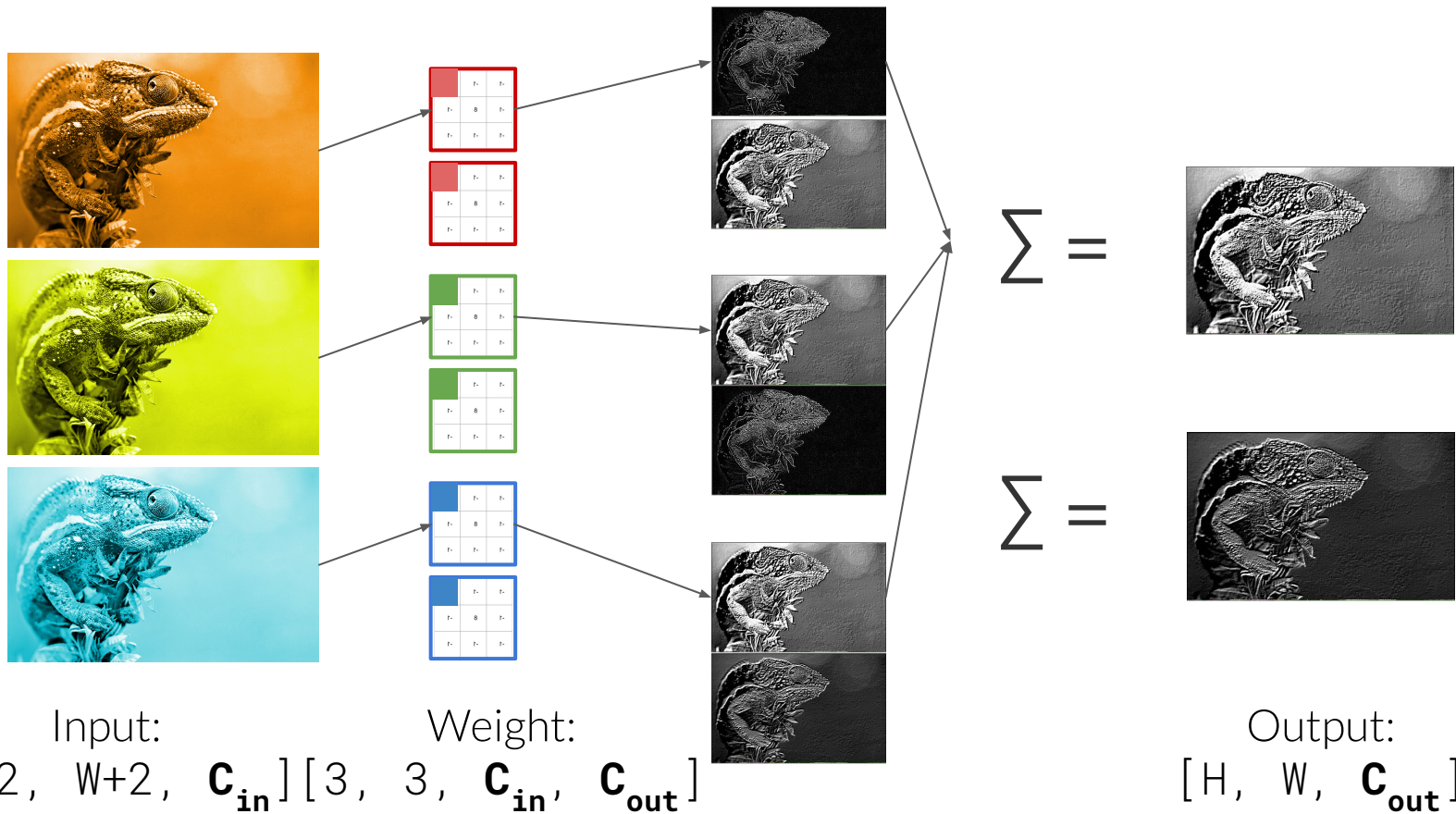
$$W^{180^\circ}$$
$$[3, 3, \mathbf{c}_{\text{out}}]$$

*



$$\delta^l$$
$$[H, W, \mathbf{c}_{\text{out}}]$$

General convolution - FORWARD



General convolution - BACKWARD



$= \Sigma$



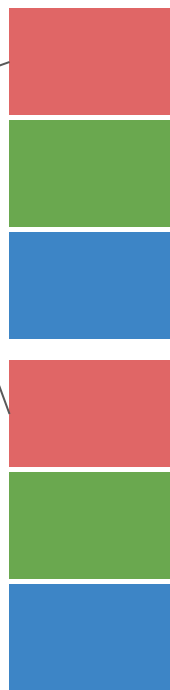
$= \Sigma$



$= \Sigma$

δ^{l-1}

$[H+2, W+2, \mathbf{C}_{in}]$

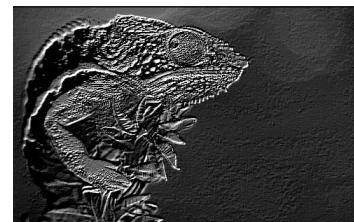
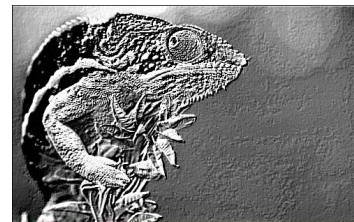


$*$

$*$

W^{180°

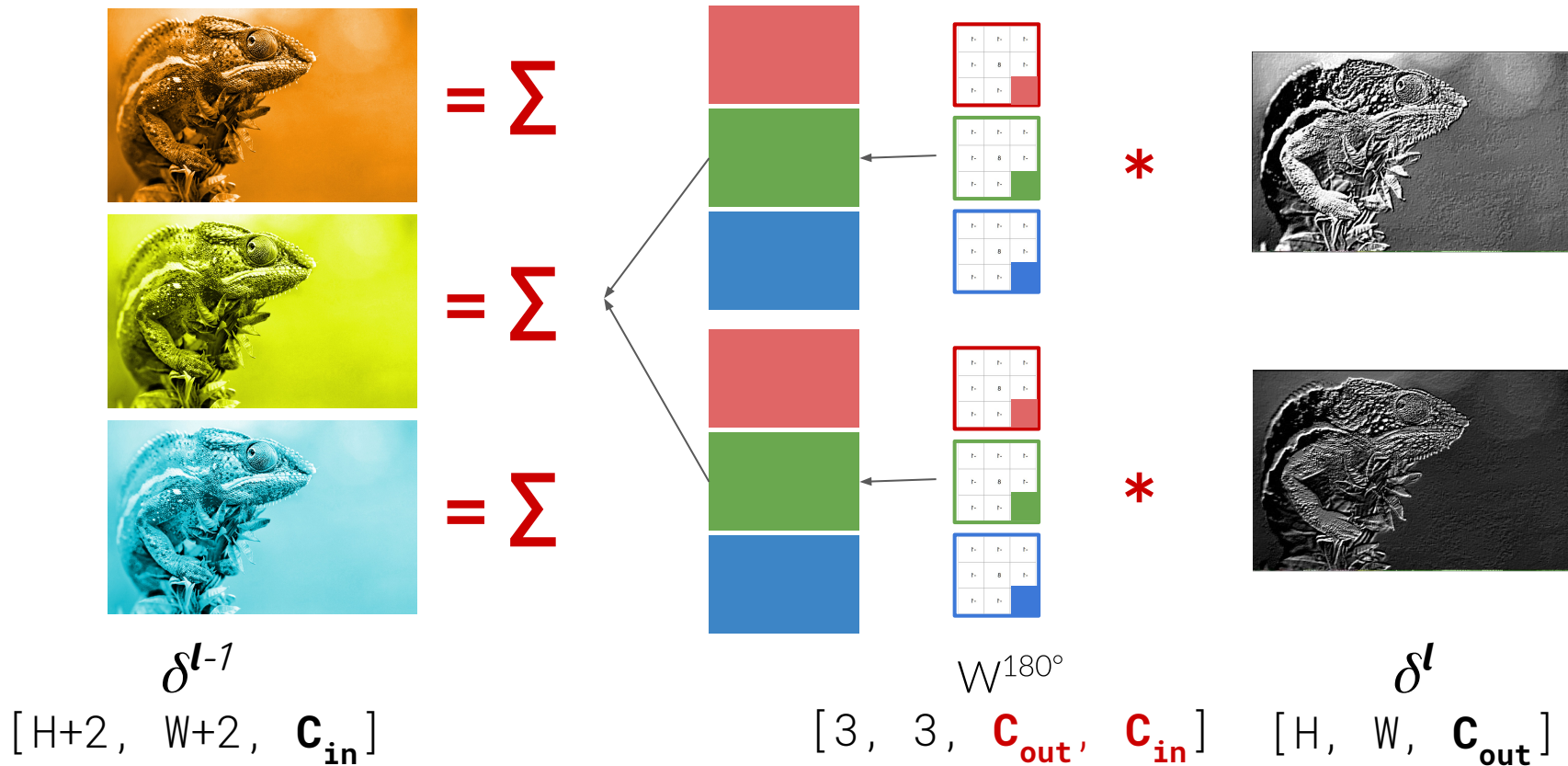
$[3, 3, \mathbf{C}_{out}, \mathbf{C}_{in}]$



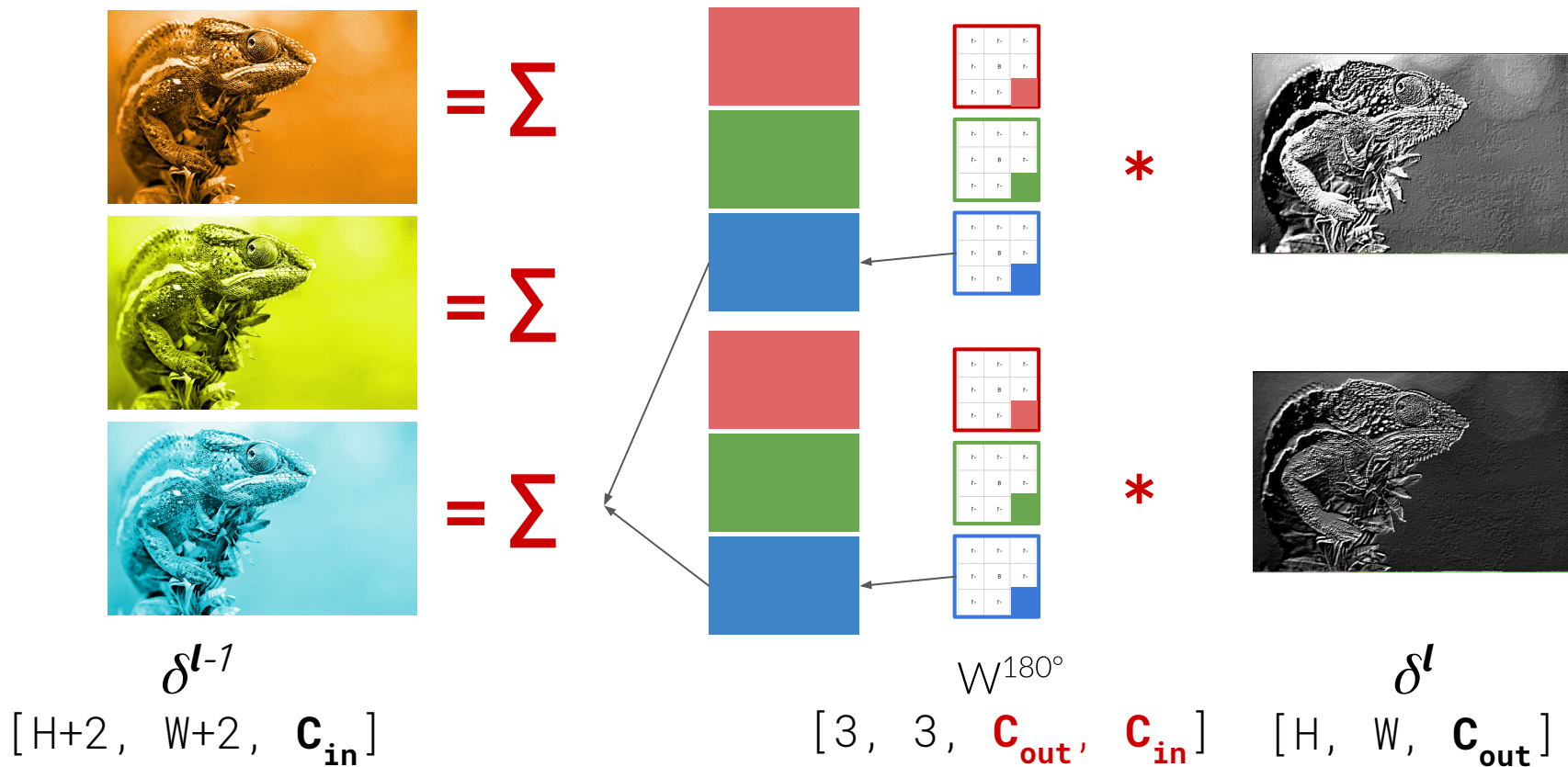
δ^l

$[H, W, \mathbf{C}_{out}]$

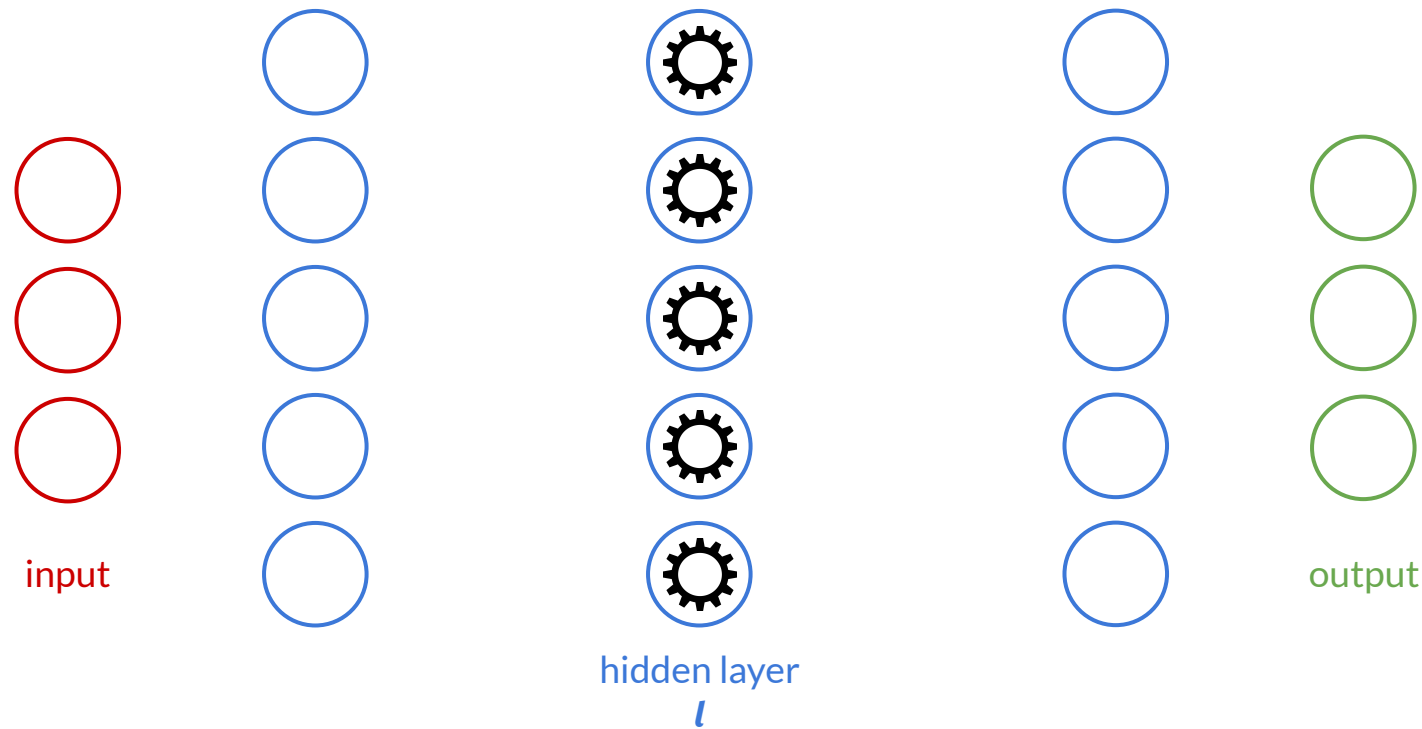
General convolution - BACKWARD



General convolution - BACKWARD



New layer's definition



1. Forward
→

2. Backward
←

3. Update 

Single convolution - update

$$\frac{dL}{dw} = \frac{\partial y}{\partial w} \delta^l$$

Dimensions?

Jacobian: $[H \times W \times C_{in}, \quad 3 \times 3 \times C_{in} \times C_{out}]$

Single convolution - update

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

$$\frac{\partial y}{\partial w} = ?$$

Single convolution - update

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

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

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

Single convolution - update

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

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

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

Single convolution - update

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

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

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

Single convolution - update

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

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

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

Single convolution - update

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

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

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

Single convolution - update

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

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

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

Single convolution - update

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

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

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

Single convolution - update

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

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

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

Single convolution - update

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

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

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

Single convolution - update

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

$$\frac{\partial y}{\partial w}$$

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

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

$$\delta^l$$

Single convolution - update

$$\frac{dL}{dw} = x^l * \delta^l$$

Dimensions?

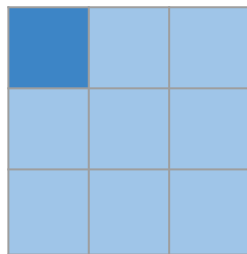
Convolution: $[H + 2, W + 2] * [H, W] = [3, 3]$

Multiple convolution - FORWARD

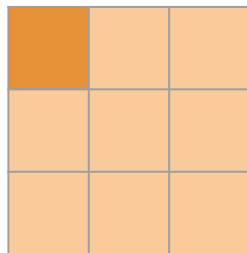


Input:
 $[H+2, W+2]$

*



=



W
 $[3, 3, \mathbf{C}_{out}]$



Output:
 $[H, W, \mathbf{C}_{out}]$

Multiple convolution - UPDATE

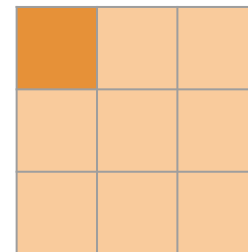
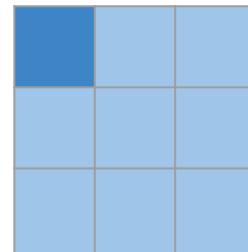


Input:
 $[H+2, W+2]$

*



=



δ^l
 $[H, W, \mathbf{C}_{\text{out}}]$

$\frac{dL}{dw}$

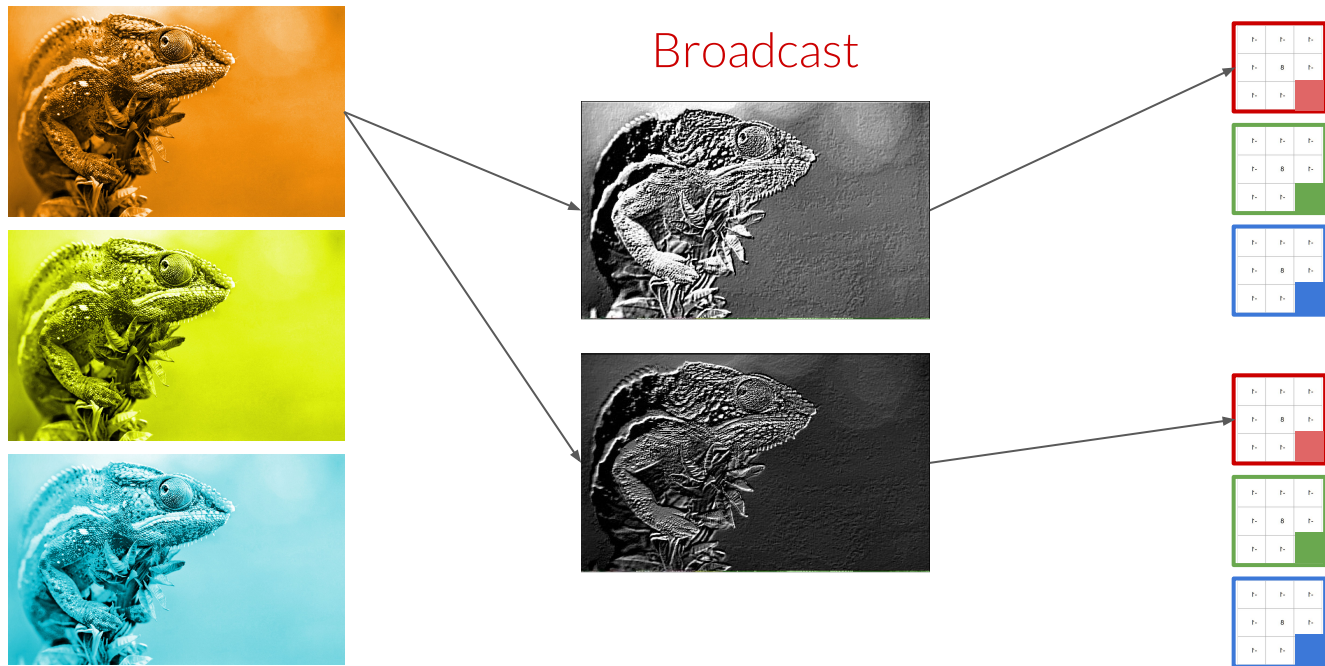
Multiple convolution - update

$$\frac{dL}{dw} = x^l * \delta^l$$

Dimensions?

Convolution: $[H + 2, W + 2] * [H, W, \mathbf{C}_{\text{out}}] = [3, 3, \mathbf{C}_{\text{out}}]$

General convolution - UPDATE

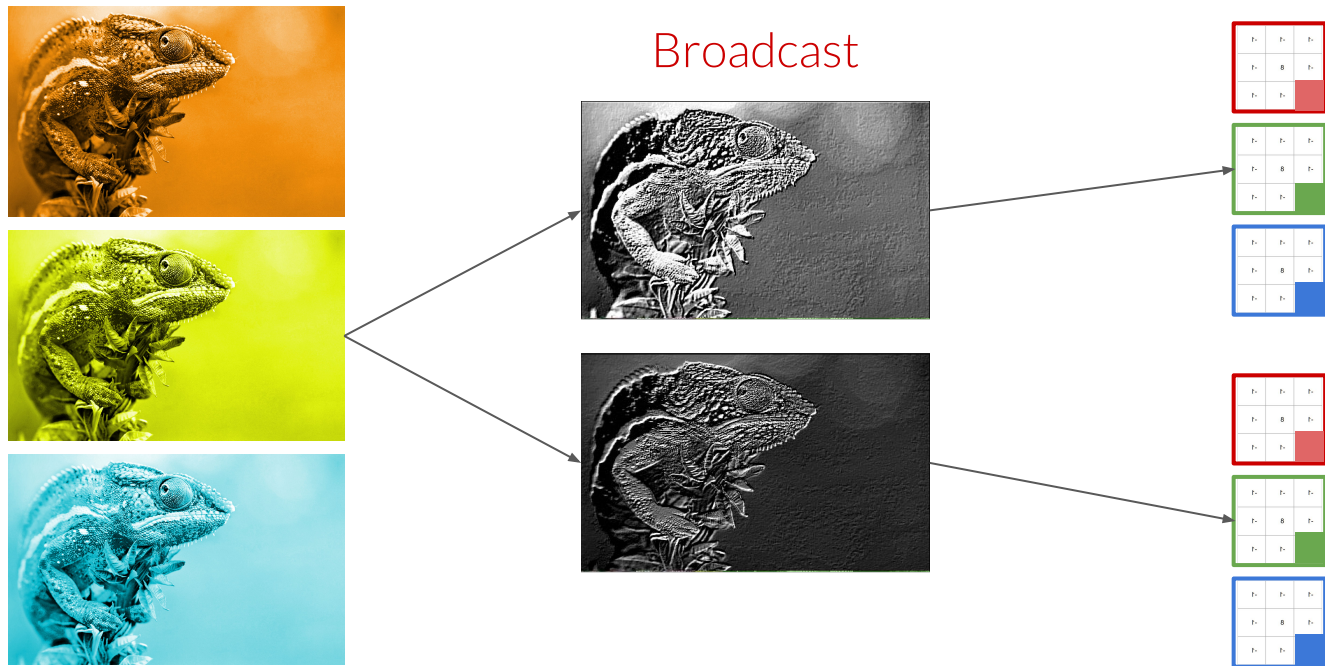


Input:
 $[H+2, W+2, \mathbf{C}_{in}]$

δ^l
 $[H, W, \mathbf{C}_{in}, \mathbf{C}_{out}]$

∇_w^L
 $[3, 3, \mathbf{C}_{in}, \mathbf{C}_{out}]$

General convolution - UPDATE



Input:
 $[H+2, W+2, \mathbf{C}_{in}]$

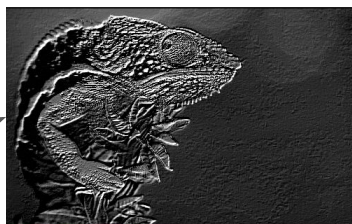
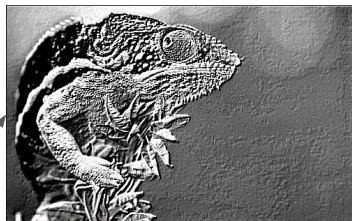
δ^l
 $[H, W, \mathbf{C}_{in}, \mathbf{C}_{out}]$

∇_w^L
 $[3, 3, \mathbf{C}_{in}, \mathbf{C}_{out}]$

General convolution - UPDATE



Broadcast



Input:
 $[H+2, W+2, \mathbf{C}_{in}]$

δ^l
 $[H, W, \mathbf{C}_{in}, \mathbf{C}_{out}]$

∇_w^L
 $[3, 3, \mathbf{C}_{in}, \mathbf{C}_{out}]$

Single convolution - update

$$\frac{dL}{dw} = x^l * \delta^l$$

Dimensions?

Convolution: $[H+2, W+2, \overset{\text{Broadcast}}{\mathbf{c}_{in}}] * [H, W, \mathbf{1}, \mathbf{c}_{out}] = [3, 3, \mathbf{c}_{in}, \mathbf{c}_{out}]$

Forward



X

$*$

W

$=$

Y

Backward



δ^{l-1}

$=$

W^{180°

$*$

δ^l

Update



$\nabla_w L$

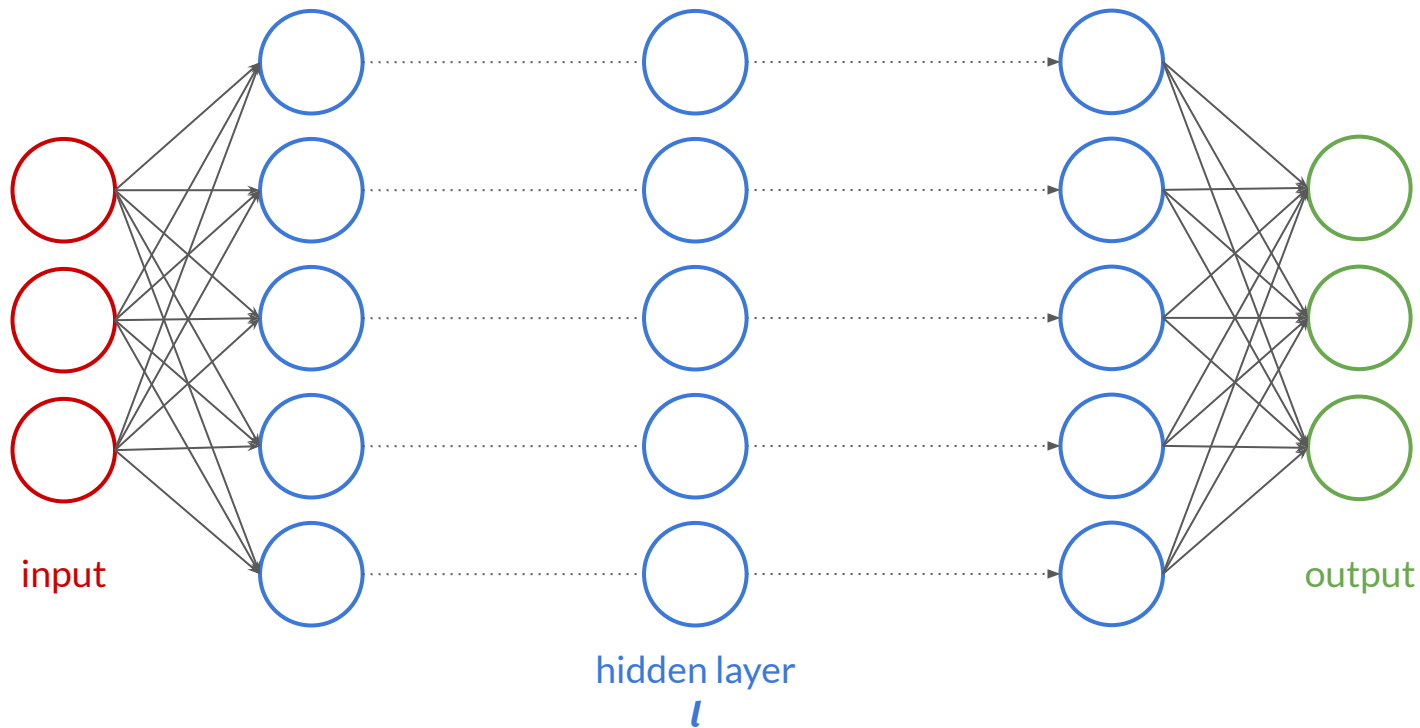
$=$

X

$*$

δ^l

New layer's definition



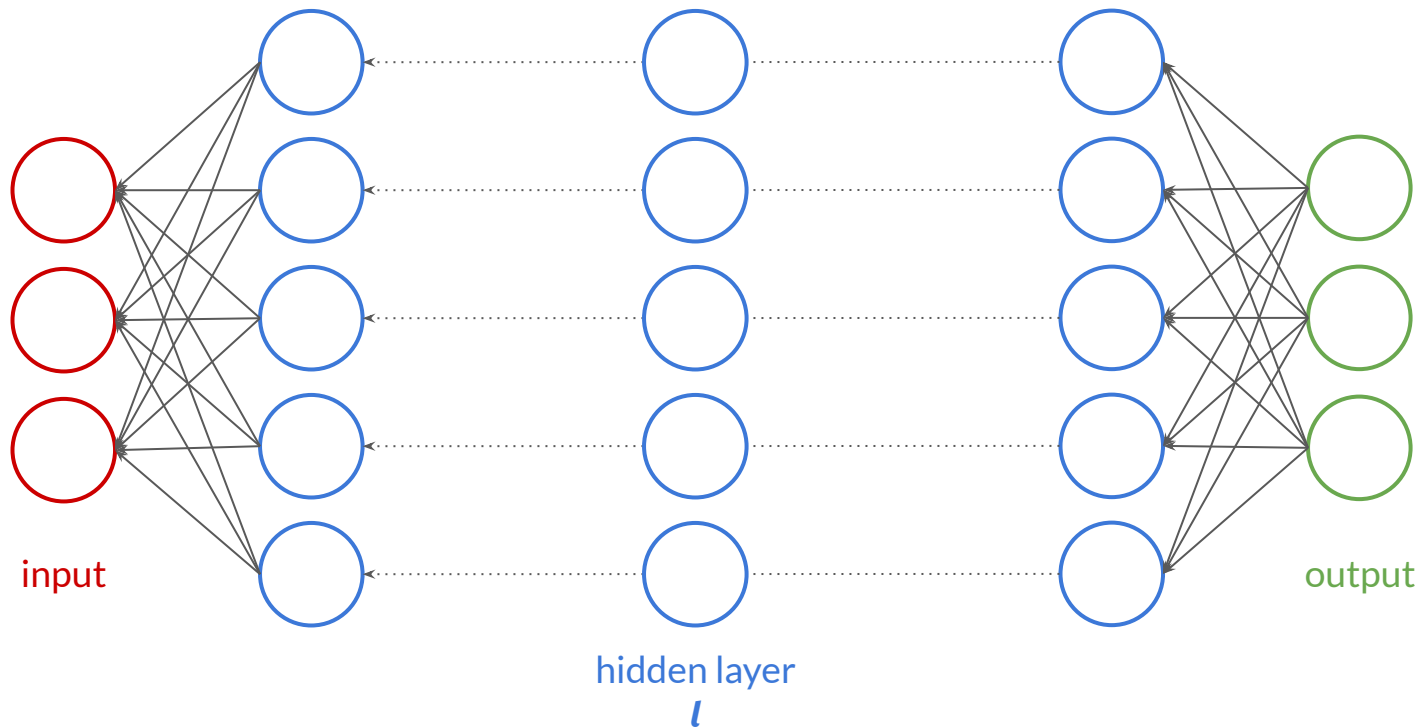
1. Forward



2. Backward

3. Update

New layer's definition

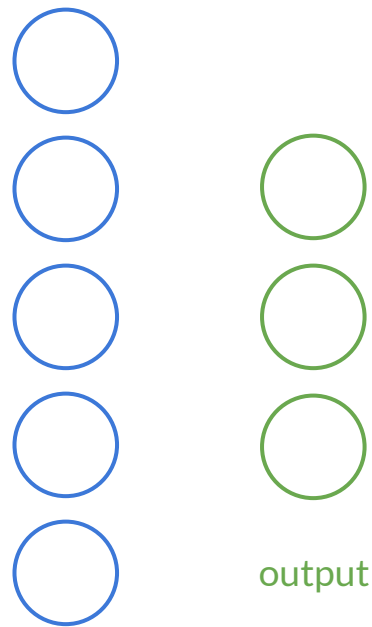
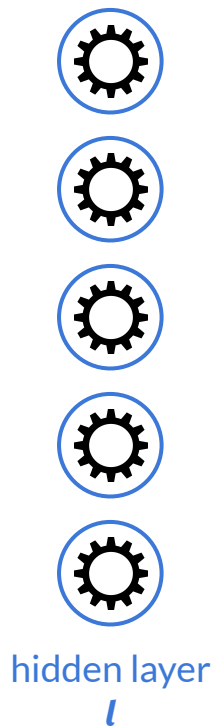
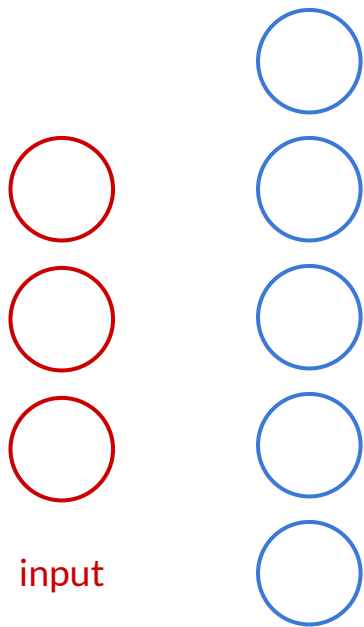


1. Forward

2. Backward

3. Update

New layer's definition



1. Forward



2. Backward



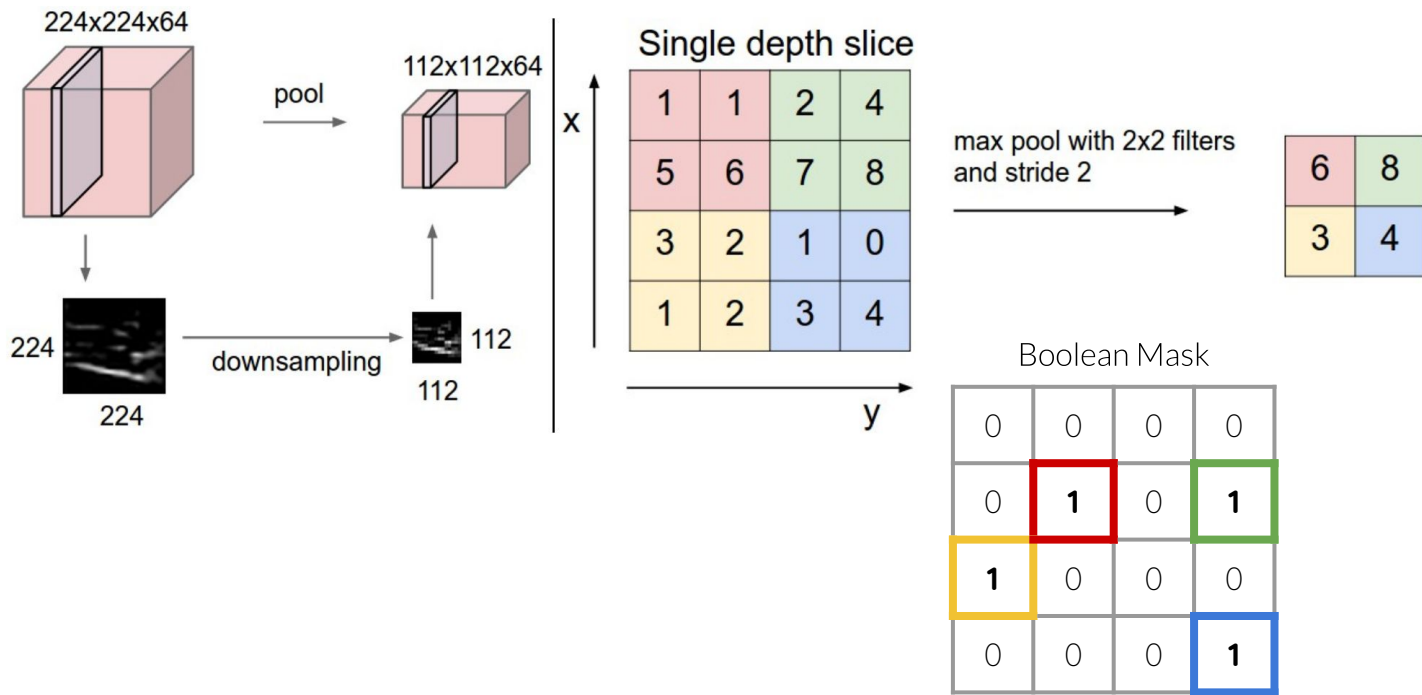
3. Update 

Max pool - Forward

1	3	2	9
7	4	1	5
8	5	2	3
4	2	1	4

7	9
8	

Max pooling - Forward



Max pooling - Forward

Single depth slice

1	1	2	4
5	6	7	8
3	2	1	0
1	2	3	4

6	8
3	4

Boolean Mask

0	0	0	0
0	1	0	1
1	0	0	0
0	0	0	1

Max pooling - Backward

Single depth error

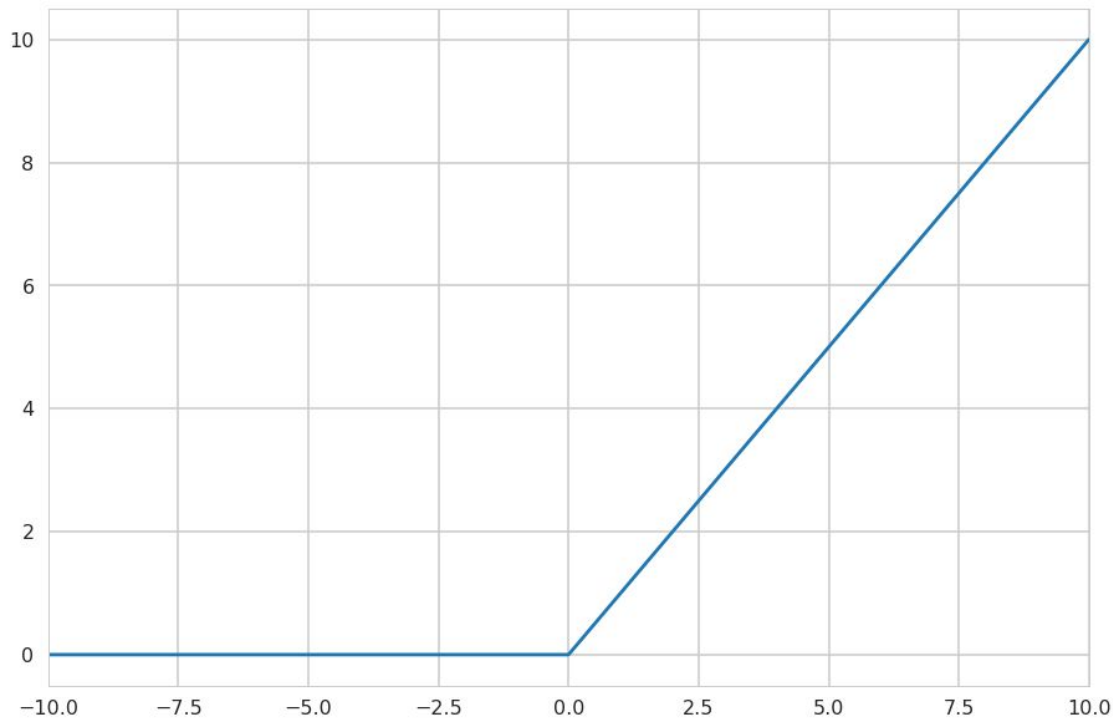
0	0	0	0
0	δ_{11}	0	δ_{12}
δ_{21}	0	0	0
0	0	0	δ_{22}

δ_{11}	δ_{12}
δ_{21}	δ_{22}

Rectified Linear Unit (ReLU)

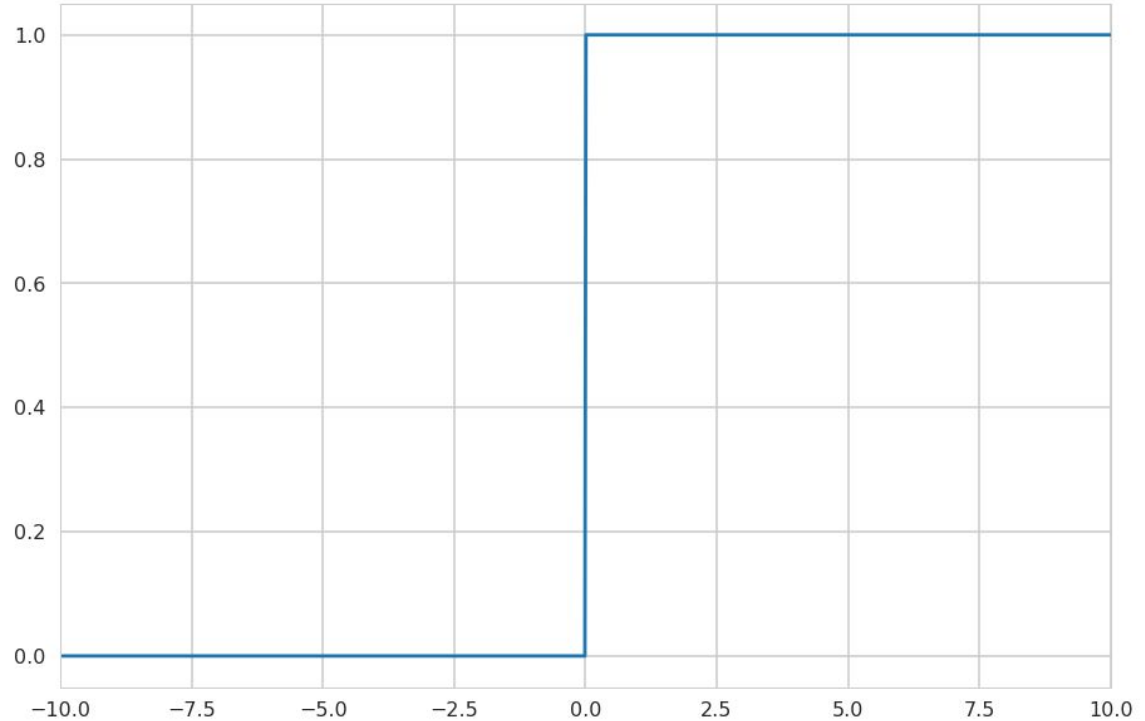
$$\text{relu}(x) = \begin{cases} x & \text{if } x > 0 \\ 0 & \text{if } x \leq 0 \end{cases}$$

2010 - Hinton et al.



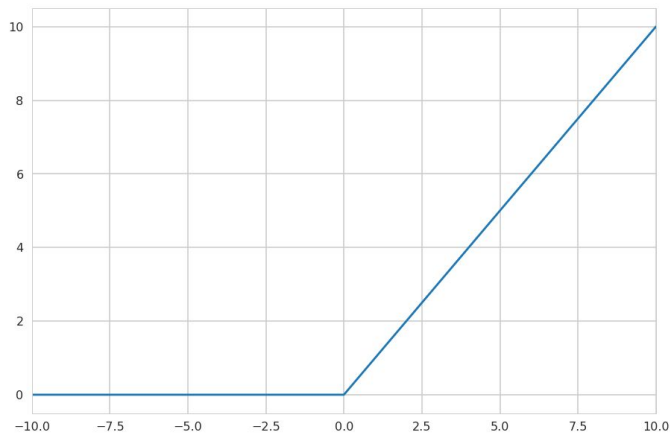
Rectified Linear Unit (ReLU) - Backward

$$\text{relu}'(x) = \begin{cases} 1 & \text{if } x > 0 \\ 0 & \text{if } x \leq 0 \end{cases}$$



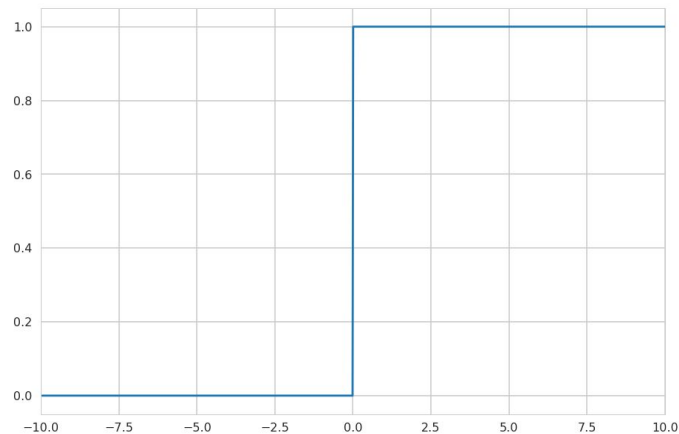
Rectified Linear Unit (ReLU) - Backward

Forward



$$\text{relu}(x) = \begin{cases} x & \text{if } x > 0 \\ 0 & \text{if } x \leq 0 \end{cases}$$

Backward



$$\text{relu}'(x) = \begin{cases} 1 & \text{if } x > 0 \\ 0 & \text{if } x \leq 0 \end{cases}$$

CIFAR 10

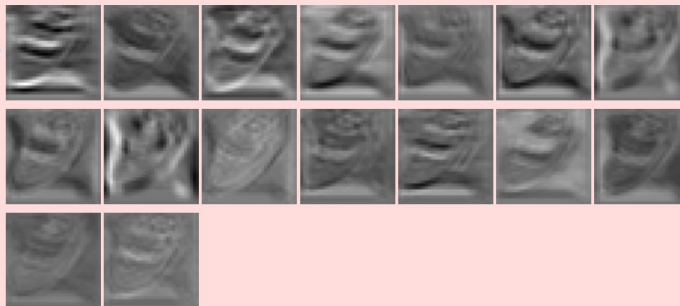
input (32x32x3)
max activation: 0.5, min: -0.49608
max gradient: 0.11953, min: -0.13089

Activations:



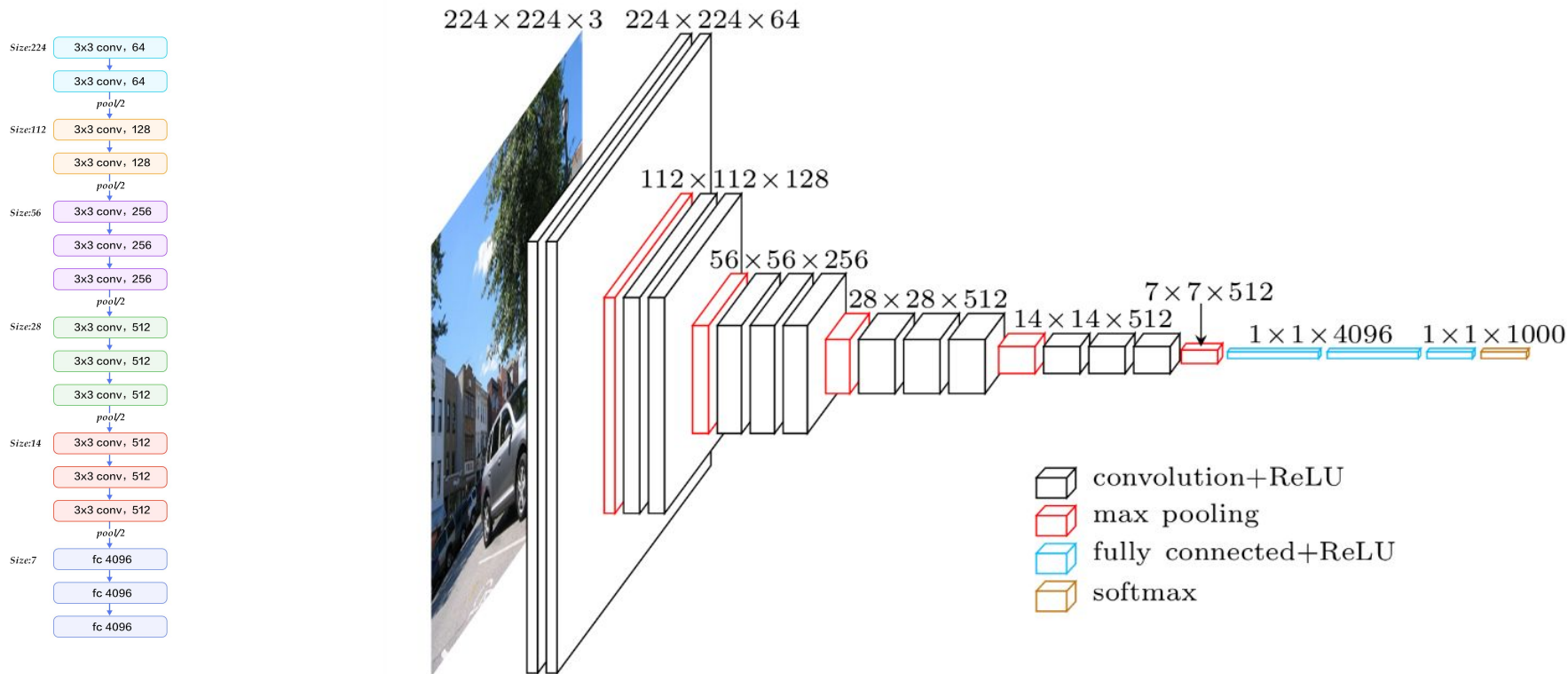
conv (32x32x16)
filter size 5x5x3, stride 1
max activation: 1.85529, min: -1.75525
max gradient: 0.09955, min: -0.11062
parameters: $16 \times 5 \times 5 \times 3 + 16 = 1216$

Activations:

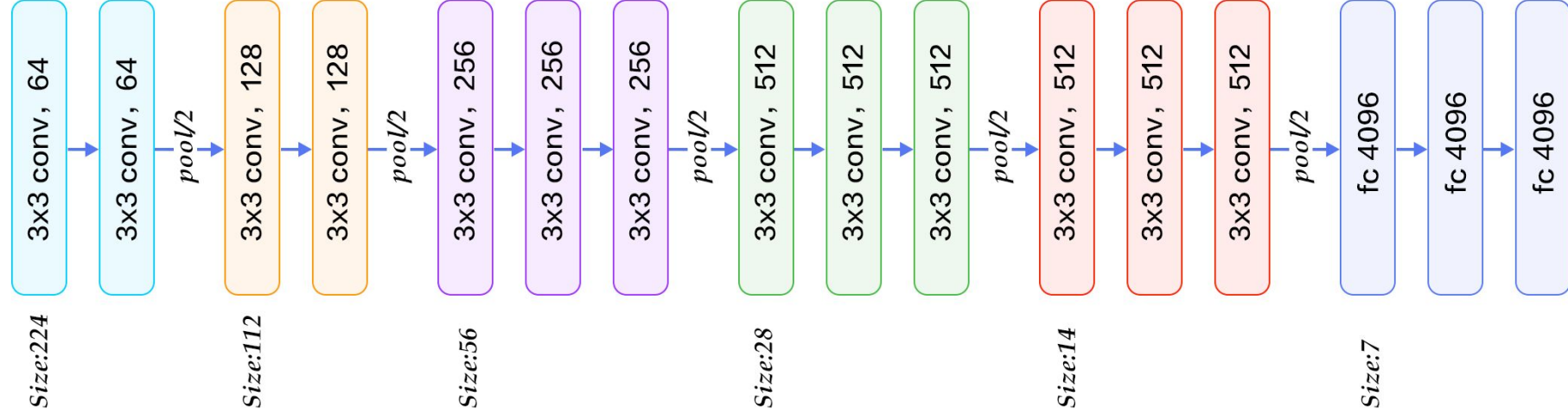


Live demo: <http://setosa.io/ev/image-kernels/>

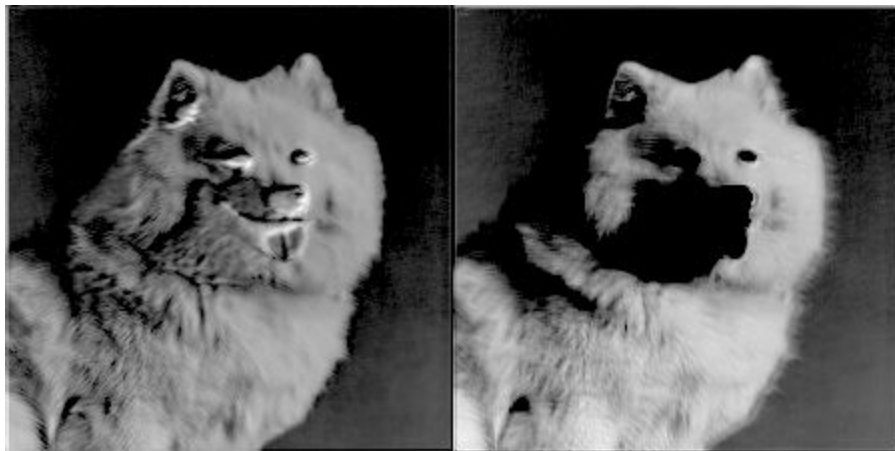
VGG 16



VGG 16



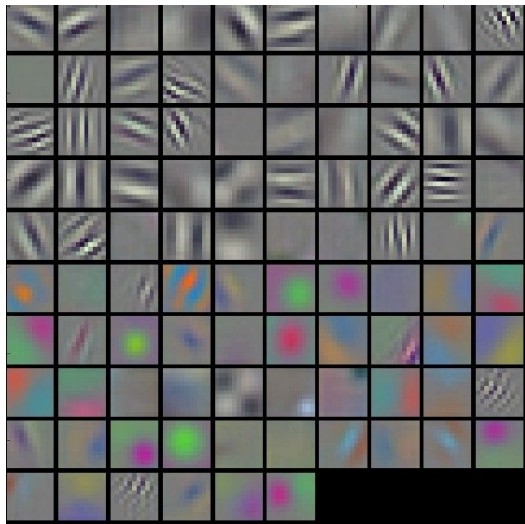
Activation maps under training



early layers -> edge detectors

Activation maps under training

WEIGHTS

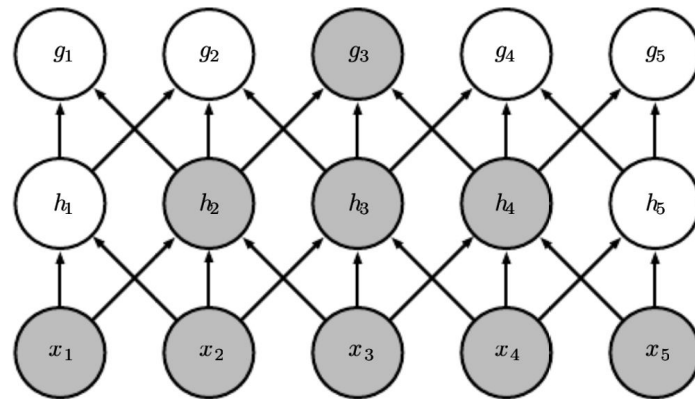
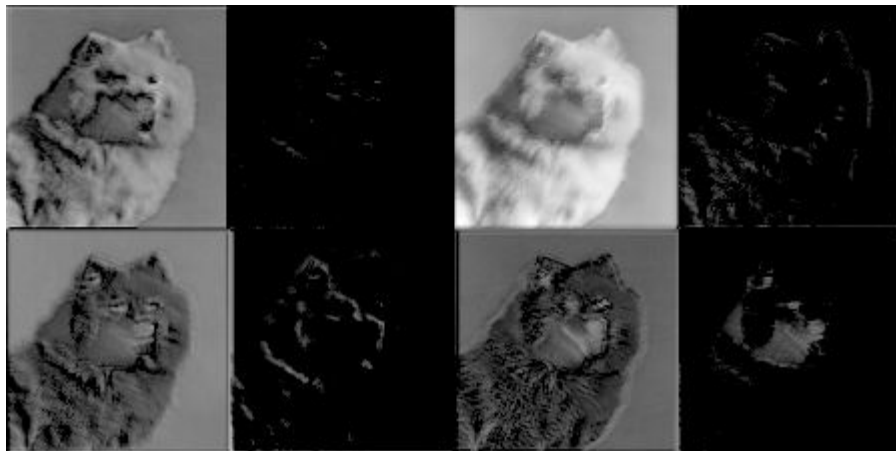


WEIGHTS



early layers -> edge detectors

Activation maps under training



latent layers -> no direct connection

Visualizing latent layers

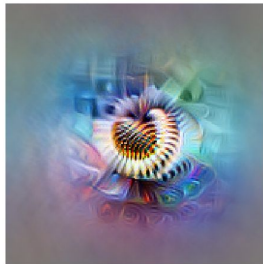
Dataset Examples

show us what neurons respond to in practice



Optimization

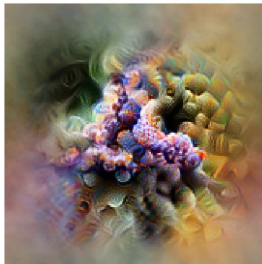
isolates the causes of behavior from mere correlations. A neuron may not be detecting what you initially thought.



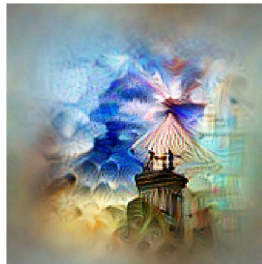
Baseball—or stripes?
mixed4a, Unit 6



Animal faces—or
snouts?
mixed4a, Unit 240



Clouds—or fluffiness?
mixed4a, Unit 453



Buildings—or sky?
mixed4a, Unit 492

Bias?

“ ... We're not using biases because meh... ”

-- Andrej Karpathy

source: <http://karpathy.github.io/2016/05/31/rl/>