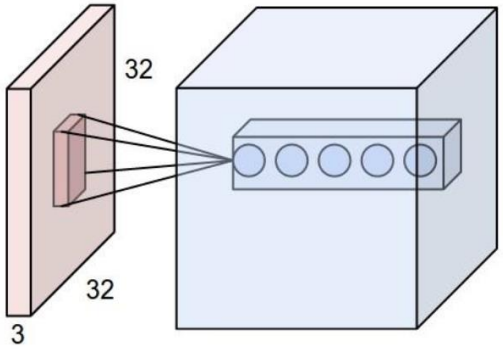
Neural Network Segmentation - Basic structure

Feature Representation by Convolution

A convolutional layer is not fully connected, but has a narrowed down receptive field (e.g. 3x3).

The parameters of each filter are spatially shared: A feature that is useful in one place, ought to be useful in another, too.

Downsampling
Upsampling
Parameter sharing



- **Depth:** number of filters
- **Stride:** filter step size (when we "slide" it)
- **Padding:** zero-pad the input

https://selfdrivingcars.mit.edu/

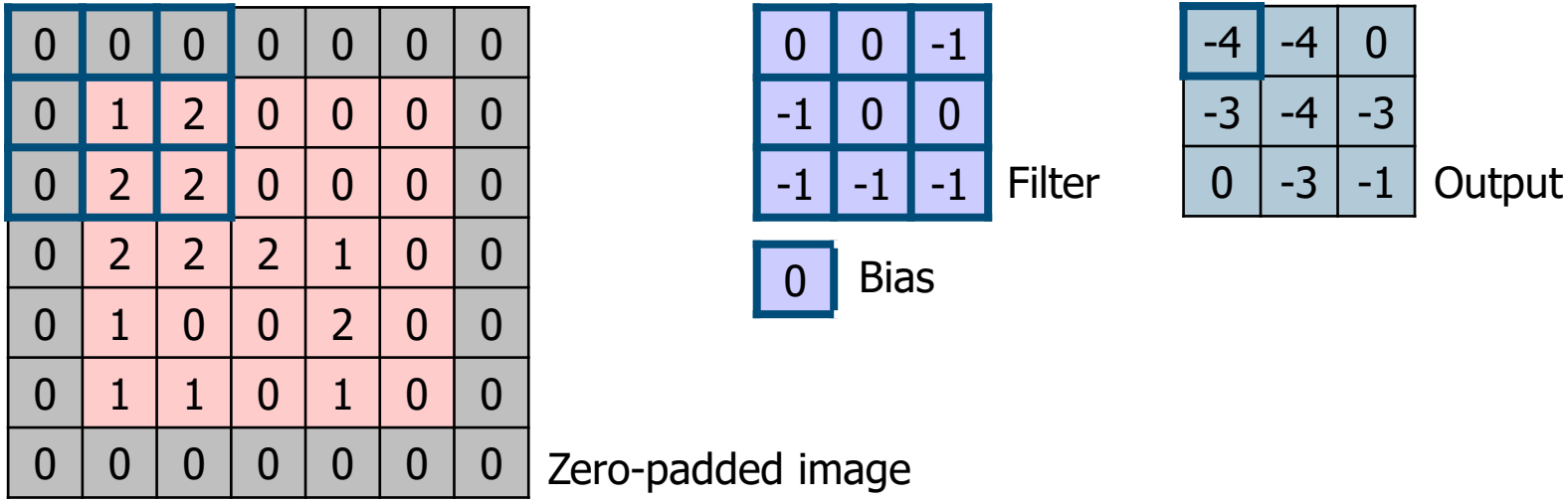
Feature Representation by Convolution

0	0	0	0	0	0	0
0	1	2	0	0	0	0
0	2	2	0	0	0	0
0	2	2	2	1	0	0
0	1	0	0	2	0	0
0	1	1	0	1	0	0
0	0	0	0	0	0	0

Zero-Padding is adding zero-valued pixel to the image border (gray area).

Downsampling
Upsampling
Parameter sharing

Feature Representation by Convolution



Downsampling

Upsampling

Parameter sharing

Feature Representation by Convolution

0	0	0	0	0	0	0
0	1	2	0	0	0	0
0	2	2	0	0	0	0
0	2	2	2	1	0	0
0	1	0	0	2	0	0
0	1	1	0	1	0	0
0	0	0	0	0	0	0

0	0	-1
-1	0	0
-1	-1	-1

Filter

0

Bias

-4	-4	0
-3	-4	-3
0	-3	-1

Output

Amount of filters or convolution depth: 1
Filter step size or Stride: 2
Zero-padded image

Downsampling
Upsampling
Parameter sharing

Feature Representation by Convolution

0	0	0	0	0	0	0
0	1	2	0	0	0	0
0	2	2	0	0	0	0
0	2	2	2	1	0	0
0	1	0	0	2	0	0
0	1	1	0	1	0	0
0	0	0	0	0	0	0

0	0	-1
-1	0	0
-1	-1	-1

Filter

0 Bias

-4	-4	0
-3	-4	-3
0	-3	-1

Output

Downsampling
Upsampling
Parameter sharing

Review edge detector:
Similar idea, now the parameters of the filters are learned. We want a lot of filters!

$$L_x = \begin{bmatrix} +1 & 0 & -1 \\ +2 & 0 & -2 \\ +1 & 0 & -1 \end{bmatrix} L \quad \text{and} \quad L_y = \begin{bmatrix} +1 & +2 & +1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} L.$$

Lena Test image

Lena , the ‘hello world!’ of image processing. 330x330

Cover photo of 1972 Playboy magazine of the Swedish model Lena Söderberg.

Since then Lena was a guest at several IEEE conferences. The image also sparked discussions on gender-equality in the male-dominated field of engineering.

It is a good test image because of its detail, flat regions, shading, and texture.

https://en.wikipedia.org/wiki/Lenna

Thresholding

Edge detection

Segment boundaries and edges are closely related.

Since there is often a large gradient at the segment boundaries.

Canny Edge Detection

$$L_x = \begin{bmatrix} +1 & 0 & -1 \\ +2 & 0 & -2 \\ +1 & 0 & -1 \end{bmatrix} L \quad \text{and} \quad L_y = \begin{bmatrix} +1 & +2 & +1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} L.$$

Sobel Operators for Edge Detection

Clustering

Region growing

Feature Representation by Convolution

0	0	0	0	0	0	0
0	1	2	0	0	0	0
0	2	2	0	0	0	0
0	2	2	2	1	0	0
0	1	0	0	2	0	0
0	1	1	0	1	0	0
0	0	0	0	0	0	0

0	0	-1
-1	0	0
-1	-1	-1

Filter

0

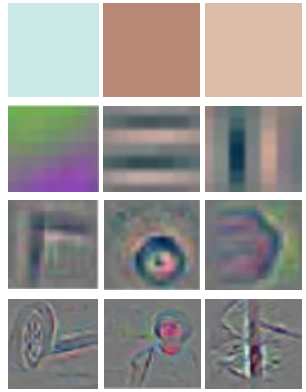
Bias

-4	-4	0
-3	-4	-3
0	-3	-1

Output

Downsampling
Upsampling
Parameter sharing

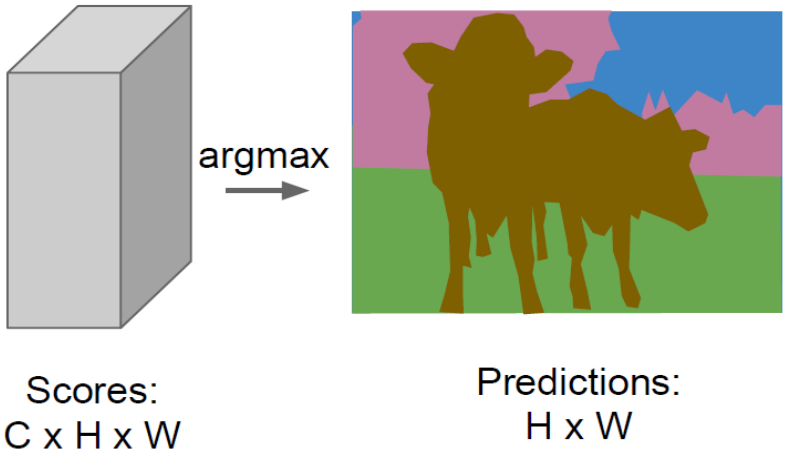
Review edge detector:
Similar idea, now the parameters of the filters are learned. And we want to go deep!



Last layer

Last Layer results in a tensor with H x W image resolution and a depth of
C: Number of classes to segment.

The last layer should encode the values into a range of values of (0;1).
Either by softmax or sigmoid function.



http://cs231n.github.io/

Cross-entropy
Dice-coefficient

Neural Network Activation Functions - Review

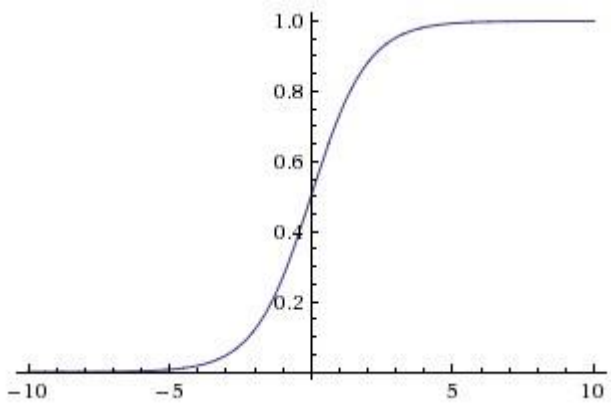
Sigmoid

$$f(z) = 1 - \frac{1}{1 + e^{-x}}$$

$$f'(z) = (1 - f(z))f(z)$$

Binary classification only.

The probability sum does not need to be one.



http://cs231n.github.io/neural-networks-1/

Softmax

Normalized exponential function

$$\sigma(\mathbf{z})_i = \frac{e^{z_i}}{\sum_{j=1}^{K} e^{z_j}}$$

Used for multi-class segmentation.

Probability sum will be 1.

before	after
2.0	0.7
1.0	0.2
0.1	0.1