

Deep Learning

Convolution Neural Network Model

Jan Platoš, Radek Svoboda

February 23, 2022

Department of Computer Science
Faculty of Electrical Engineering and Computer Science
VŠB - Technical University of Ostrava

Convolution Neural Network Model

A convolution is defined as the integral of the product of the two functions after one is reversed and shifted. It is a mathematical way how to analyze behavior of the functions and the relation between the functions.

A convolution is defined as the integral of the product of the two functions after one is reversed and shifted. It is a mathematical way how to analyze behavior of the functions and the relation between the functions.

In image processing, *kernel* or *convolution matrix* or *mask* is a small matrix. In general the convolution in image processing is defined as:

$$g(x, y) = \omega * f(x, y) = \sum_{s=-a}^a \sum_{t=-b}^b \omega(s, t) f(x - s, y - t)$$

where $g(x, y)$ is filtered image, $f(x, y)$ is original image, ω if the filter kernel.

A kernel (also called a filter) is a smaller-sized matrix in comparison to the dimensions of the input image, that consists of real valued entries.

Sample Convolution Kernels

Identity

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Sobel vertical
edge detection

$$\begin{bmatrix} +1 & 0 & -1 \\ +2 & 0 & -2 \\ +1 & 0 & -1 \end{bmatrix}$$

Sobel horizontal
edge detection

$$\begin{bmatrix} +1 & +2 & +1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}$$

Edge detection

$$\begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix}$$

Sharpen

$$\begin{bmatrix} 0 & -1 & 0 \\ -1 & 5 & -1 \\ 0 & -1 & 0 \end{bmatrix}$$

Uniform blur

$$\frac{1}{9} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

Gaussian blur 3x3

$$\frac{1}{16} \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}$$

Size of the kernel defines the dimensions of the kernels.

Number of input channels reflects the number of channels of the image (grayscale, RGB, etc.)

Number of output channels defines the number of kernels applied on the image, and, therefore, the output of the layer.

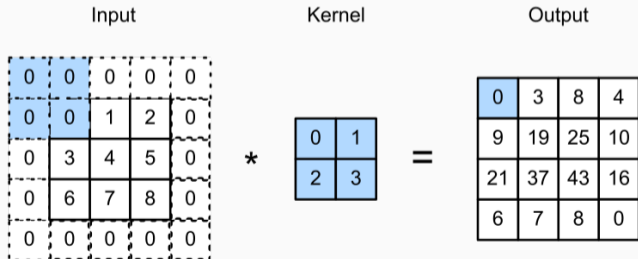
Stride is the size of the step that kernel is moved on the image.

Padding is system the kernel is moder on the image.

Weights sharing

Padding

One tricky issue when applying convolutional is losing pixels on the edges of our image. A straightforward solution to this problem is to add extra pixels around the boundary of our input image, which increases the effective size of the image.



Pooling

Pooling is a way how to decrease the amount of information transferred from one layer to another. The standard way how to do it is *Average Pooling* and *Maximum Pooling*.

4	6	1	3
0	8	12	9
2	3	16	100
1	46	74	27



8	12
46	100

(i)

35	19	25	6
13	22	16	63
4	3	7	10
9	8	1	3



35	63
9	10

(iii)

9	7	3	2
26	37	14	1
15	29	16	0
8	6	54	2



37	14
29	54

(ii)

35	19	25	6
13	22	16	63
4	3	7	10
9	8	1	3



35	25	63
22	22	63
9	8	10

(iv)

Questions?