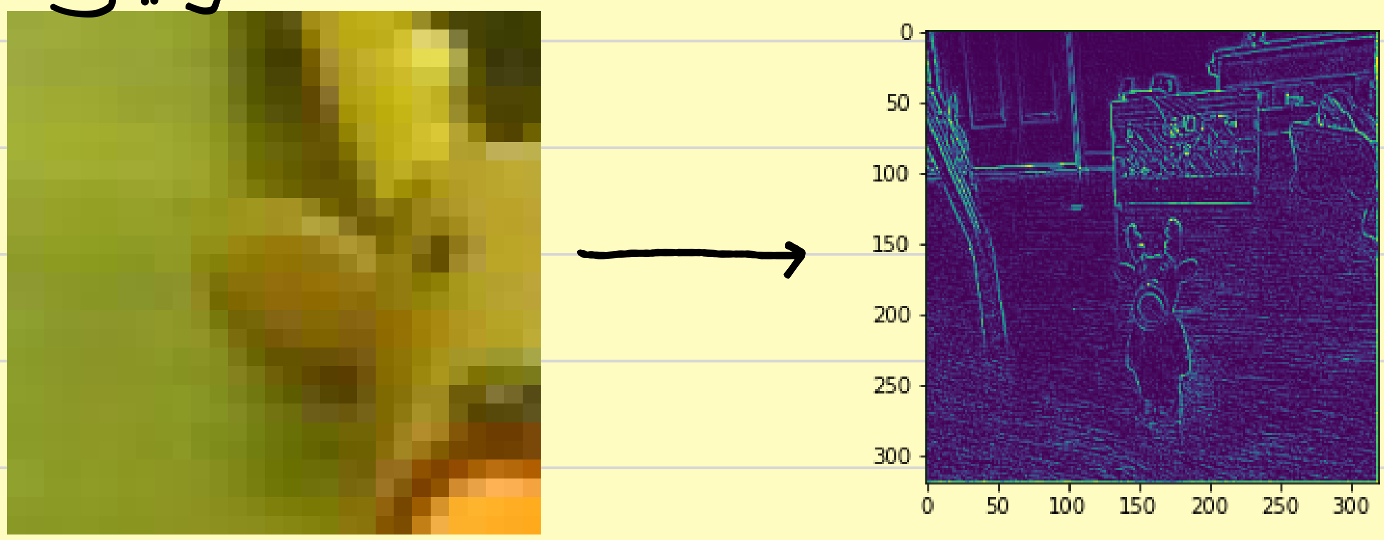
"Deep learning by hand". Part 3. CONVOLUTION.

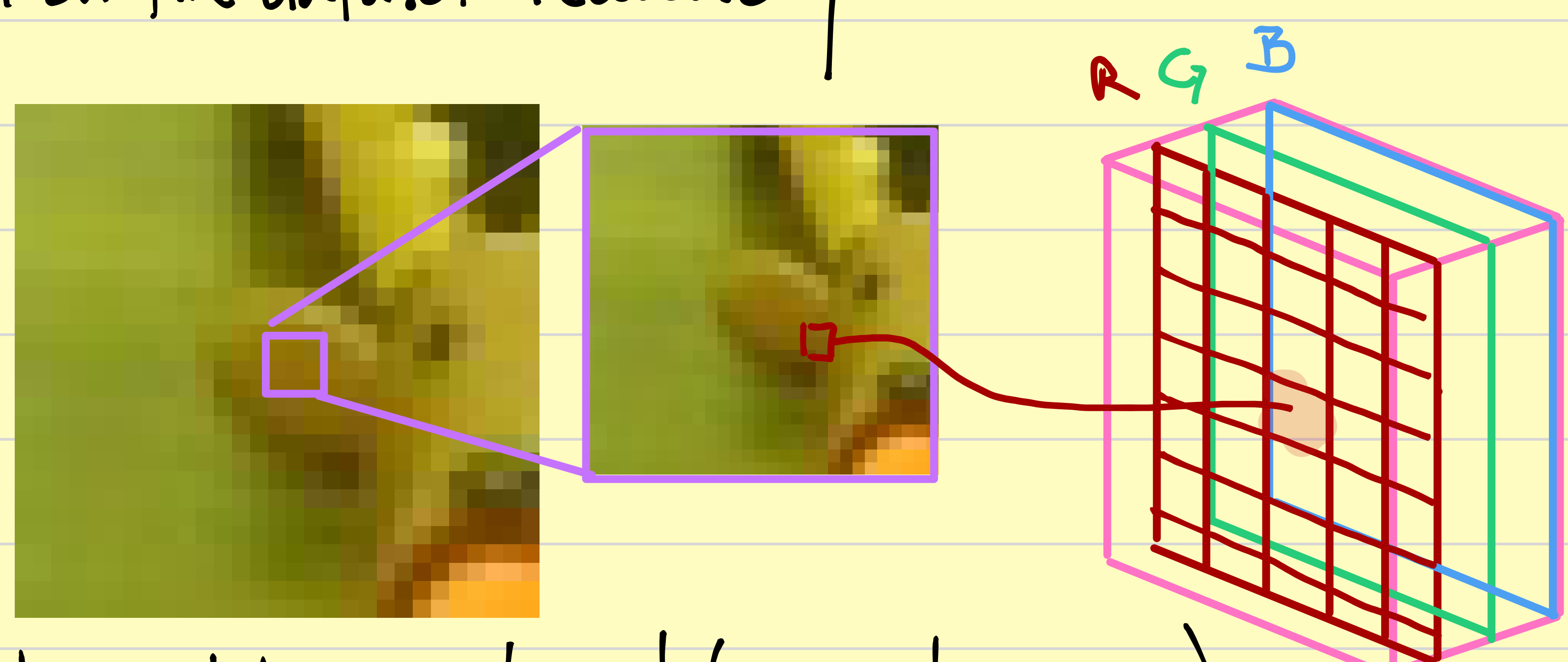


Background: convolution is a technic/method used in the datatreatment of euclidean domains (i.e. Images, Videos, ...) in order to extract patterns (i.e. edges, colors, ...) from the dataset.

Convolution uses "filters" to extract the information/patterns. These filters are called "KERNELS".

A kernel, in this context, is a tensor of "weights" that is used on the dataset recursively.

Color Image



Each pixel has 3 channels (in a color image).

These channels can attain 256 different value.

Each channel [R, G, B] has 1 byte for each pixel.
1 byte = 8 bits, and each bit can have two values [0 or 1].

0	0	0	0	0	0	0	0

$2^8 = 256$

In an euclidean dataset, we can meassure distance.

$\boxed{\text{INPUT KERNEL or FILTER.}}$ $\boxed{(f \circ g)(x) = f(g(x))}$

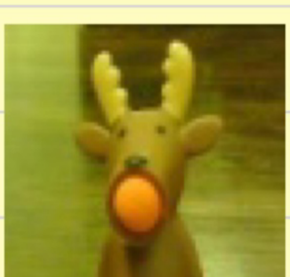
The kernel has several properties:

1) Kernels can have different shapes.
2) Kernels filter information by adapting the inpt euklydean dataset.

This gives us diffent types of kernels:

- edge detection → 3x3 Edge detect $\begin{bmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{bmatrix}$
- blurring
- sharpening
- attention → 3x3 Attention $\begin{bmatrix} 0 & -1 & 0 \\ -1 & 8 & -1 \\ 0 & -1 & 0 \end{bmatrix}$

Beispiel:

 →

R

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

G

B

Attention kernel: $\begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{bmatrix}$

STRIDE: how many pixels does the kernel "move"

STRIDE: 1 pixel

R

$$\begin{bmatrix} 1 & 1 & 0 & 2 \\ 1 & 2 & 1 & 1 \\ 0 & 1 & 2 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{bmatrix} = 1.0 + 1.(-1) + 0.0 + 1.(-1) + 2.4 + 1.(-1) + 0.0 + 1(-1) + 2.0 = 4$$

R → 1 pixel

$$\begin{bmatrix} 1 & 1 & 0 & 2 \\ 1 & 2 & 1 & 1 \\ 0 & 1 & 2 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{bmatrix} = 1.0 + 0.(-1) + 2.0 + 2.(-1) + 1.4 + 1.(-1) + 1.0 + 2.(-1) + 1.0 = -1$$

1 pixel ↓

R

$$\begin{bmatrix} 1 & 1 & 0 & 2 \\ 1 & 2 & 1 & 1 \\ 0 & 1 & 2 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{bmatrix} = 1.0 + 2.(-1) + 1.0 + 0.(-1) + 1.4 + 2.(-1) + 1.0 + 0.(-1) + 0.0 = 0$$

R

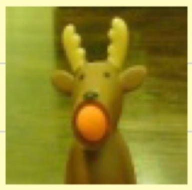
$$\begin{bmatrix} 1 & 1 & 0 & 2 \\ 1 & 2 & 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \end{bmatrix} = 2.0 + 1.(-1) + 1.0 + 1.(-1) + 4.2 + (-1).1 + 0.0 + 0.(-1) + 0.1 = 7$$

$$\begin{bmatrix} 4 & -1 \\ 0 & 7 \end{bmatrix}$$

0	1	2	1
1	0	0	1

0	-1	0

Beispiel:



R

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

G

B

Att. Filter 3x3

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

R

4	-1
0	7

G

B

Now we try a 2x2 & STRIDE=1

$$\begin{matrix} \boxed{\begin{matrix} 1 & 1 \\ 1 & 2 \end{matrix}} & \begin{matrix} 0 & 2 \\ 1 & 1 \end{matrix} \\ \begin{matrix} 0 & 1 \\ 1 & 0 \end{matrix} & \begin{matrix} 2 & 1 \\ 0 & 1 \end{matrix} \end{matrix} \cdot \begin{bmatrix} 2 & 0 \\ -1 & 1 \end{bmatrix}$$

$2 \cdot 1 + 0 \cdot 1 +$
$= (-1) \cdot 1 + 1 \cdot 2 = 3$

$$\begin{matrix} 1 & \boxed{\begin{matrix} 1 & 0 \\ 2 & 1 \end{matrix}} & 2 \\ 1 & & 1 \\ 0 & 1 \;\; 2 & 1 \\ 1 & 0 \;\; 0 & 1 \end{matrix} \cdot \begin{bmatrix} 2 & 0 \\ -1 & 1 \end{bmatrix}$$

$2 \cdot 1 + 0 \cdot 0 +$
$= (-1) \cdot 2 + 1 \cdot 1 = 1$

$$\begin{matrix} 1 \;\; 1 & \boxed{\begin{matrix} 0 & 2 \\ 1 & 1 \end{matrix}} \\ 1 \;\; 2 & \\ 0 \;\; 1 & 2 \;\; 1 \\ 1 \;\; 0 & 0 \;\; 1 \end{matrix} \cdot \begin{bmatrix} 2 & 0 \\ -1 & 1 \end{bmatrix}$$

$2 \cdot 0 + 2 \cdot 0$
$= (-1) \cdot 1 + 1 \cdot 1 = 0$

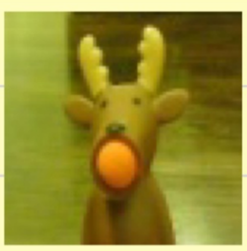
3	1	0
3	5	1
-1	2	5

$$R = \begin{bmatrix} 1 & 1 & 0 & 2 \\ 1 & 2 & 1 & 1 \\ 0 & 1 & 2 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 2 & 0 \\ -1 & 1 \end{bmatrix} \quad \begin{aligned} &2\cdot 1+0\cdot 2 \\ &+(-1)\cdot 0+1\cdot 1 = 3 \end{aligned}$$

...

PADDING.

Beispiel:



$\rightarrow$

$$\begin{array}{ccc} R & G & B \\ \begin{bmatrix} 1 & 1 & 0 & 2 \\ 1 & 2 & 1 & 1 \\ 0 & 1 & 2 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix} & \begin{bmatrix} \; \end{bmatrix} & \begin{bmatrix} \; \end{bmatrix} \end{array}$$

$\bullet$

3x3 Att. kernel $\begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{bmatrix}$

we might loose Information in the edges of an Image with certain Kernels (i.e. Attention)

Padding is about adding zeros around the image. Padding $p=1$ means we add a line of zeros around the image.

This allows us to keep all the information of the dataset after convolution.

R

$$\begin{bmatrix} 1 & 1 & 0 & 2 \\ 1 & 2 & 1 & 1 \\ 0 & 1 & 2 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix} \xrightarrow{p=1} \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 2 & 0 \\ 0 & 1 & 2 & 1 & 1 & 0 \\ 0 & 0 & 1 & 2 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Now we perform convolution with STRIDE = 2 PADDING = 1

$$\begin{matrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 2 & 0 \\ 0 & 1 & 2 & 1 & 1 & 0 \\ 0 & 0 & 1 & 2 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{matrix} \cdot \begin{matrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{matrix} = 4.1+(-1).1+1.(-1)+2.0 = 2$$

2 pixels!

$$\begin{matrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 2 & 0 \\ 0 & 1 & 2 & 1 & 1 & 0 \\ 0 & 0 & 1 & 2 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{matrix} \cdot \begin{matrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{matrix} = 1.(-1)+4.0+2.(-1)+0.2+(-1).1+1.0 = -4$$

$$\begin{matrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 2 & 0 \\ 0 & 1 & 2 & 1 & 1 & 0 \\ 0 & 0 & 1 & 2 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{matrix} \cdot \begin{matrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{matrix} = 0.0+1.(-1)+2.0+0.(-1)+0.4+1.(-1)+0.0+1.(-1)+0.0 = -2$$

2	-4
-2	7

PADDING

R

$$
\begin{bmatrix}
0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 1 & 0 & 2 & 0 \\
0 & 1 & \boxed{2} & 1 & 1 & 0 \\
0 & 0 & 1 & 2 & 1 & 0 \\
0 & 1 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}
\cdot
\begin{bmatrix}
0 & -1 & 0 \\
-1 & 4 & -1 \\
0 & -1 & 0
\end{bmatrix}
$$

$$
\begin{aligned}
&2\cdot 0 + 1\cdot(-1) + 1\cdot 0 + \\
= \; &1\cdot(-1) + 2\cdot 4 + 1\cdot(-1) + \\
&0\cdot 0 + 0\cdot(-1) + 1\cdot 0 = 7
\end{aligned}
$$