Convolution

J Junaid Mahmud

- We know about Convolutional Neural Networks (CNN). Here we will learn about convolution.

Polynomials

$$(2 + 3x + 4x^2)$$
$$(2 + 3x + 4x^2 + 5x^2 + 6x^4)$$

- Coefficients in polynomials.
- How to find them

$$(K_0 + K_1 + K_2x^2)$$
$$(a_0 + a_1x + a_2x^2 + a_3x^3 + a_4x^4 + a_5x^5)$$

- To find the coefficients, we should not multiply all of them. We should progress systematically.

$$(K_2a_2 + K_1a_3 + K_0a_4)x^4$$

What will be the coefficient of the term?

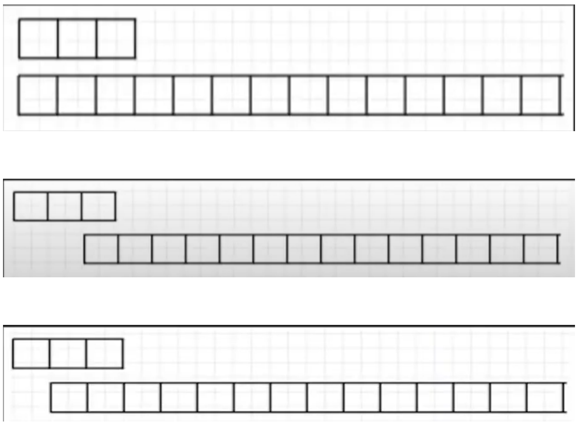
$$(K_2a_2 + K_1a_3 + K_0a_4)x^4$$

$$\begin{aligned}
(K_0a_0) \quad &+ \\
(K_1a_0 + K_0a_1)\, x \quad &+ \\
(K_2a_0 + K_1a_1 + K_0a_2)\, x^2 \quad &+ \\
(K_2a_1 + K_1a_2 + K_0a_3)\, x^3 \quad &+ \\
(K_2a_2 + K_1a_3 + K_0a_4)\, x^4 \quad &+ \\
(K_2a_3 + K_1a_4 + K_0a_5)\, x^5 \quad &+ \\
(K_2a_4 + K_1a_5)\, x^6 \quad &+ \\
(K_2a_5)\, x^7 \quad &
\end{aligned}$$

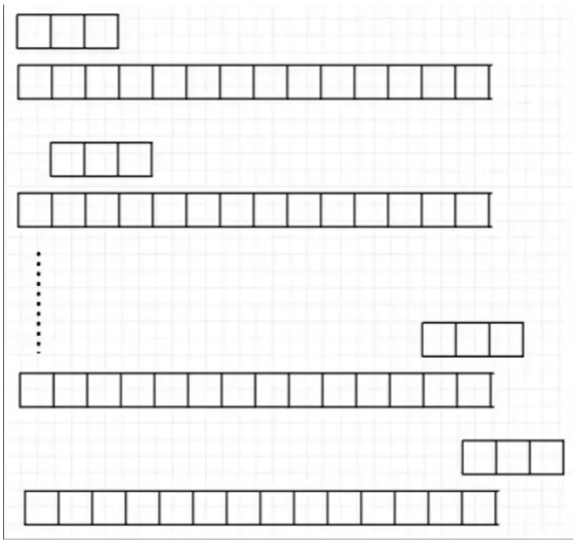
- Iterations

$$(K_0 + K_1x + K_2x^2)$$
$$(a_0 + a_1x + a_2x^2 + a_3x^3 + a_4x^4 + a_5x^5)$$

$$(K_2x^2 + K_1x + K_0)$$
$$(a_0 + a_1x + a_2x^2 + a_3x^3 + a_4x^4 + a_5x^5)$$



The first polynomial is moving with each iteration



$$(K_2x^2 + K_1x + K_0)$$
$$(a_0 + a_1x + a_2x^2 + a_3x^3 + a_4x^4 + a_5x^5)$$

$$(K_2x^2 + K_1x + K_0)$$
$$(a_0 + a_1x + a_2x^2 + a_3x^3 + a_4x^4 + a_5x^5)$$

$$(K_2x^2 + K_1x + K_0)$$
$$(a_0 + a_1x + a_2x^2 + a_3x^3 + a_4x^4 + a_5x^5)$$

$$(K_2x^2 + K_1x + K_0)$$
$$(a_0 + a_1x + a_2x^2 + a_3x^3 + a_4x^4 + a_5x^5)$$

$$\begin{aligned}
(K_0a_0) &+ \\
(K_1a_0 + K_0a_1)\,x &+ \\
(K_2a_0 + K_1a_1 + K_0a_2)\,x^2 &+ \\
(K_2a_1 + K_1a_2 + K_0a_3)\,x^3 &+ \\
(K_2a_2 + K_1a_3 + K_0a_4)\,x^4 &+ \\
(K_2a_3 + K_1a_4 + K_0a_5)\,x^5 &+ \\
(K_2a_4 + K_1a_5)\,x^6 &+ \\
(K_2a_5)\,x^7 &
\end{aligned}$$

- This type of sliding operation is called convolution. Each term of one polynomial is getting convolved with another term.

Convolution in one dimension

Convolution in 2D

Convolution in 3D

Convolution in one dimension

$$a = \left[\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right]$$
$$x = [1, 2, 3, 1, 2, 3, 1, 2, 3]$$

$$y(1) = \left\{\begin{array}{ccccccccccc} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & & & & & & & & \\ \times & \times & \times & & & & & & & & \\ 0 & 1 & 2 & 3 & 1 & 2 & 3 & 1 & 2 & 3 & 0 \end{array}\right\} = 1$$

$$y(2) = \left\{\begin{array}{ccccccccccc} & & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & & & & & & \\ & & \times & \times & \times & & & & & & \\ 0 & 1 & 2 & 3 & 1 & 2 & 3 & 1 & 2 & 3 & 0 \end{array}\right\} = 2$$

$$y(3) = \left\{\begin{array}{ccccccccccc} & & & & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & & & & \\ & & & & \times & \times & \times & & & & \\ 0 & 1 & 2 & 3 & 1 & 2 & 3 & 1 & 2 & 3 & 0 \end{array}\right\} = 2$$

$$y(4) = \left\{\begin{array}{ccccccccccc} & & & & & & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & & \\ & & & & & & \times & \times & \times & & \\ 0 & 1 & 2 & 3 & 1 & 2 & 3 & 1 & 2 & 3 & 0 \end{array}\right\} = 2$$

$$y(5) = \left\{\begin{array}{ccccccccccc} & & & & & & & & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ & & & & & & & & \times & \times & \times \\ 0 & 1 & 2 & 3 & 1 & 2 & 3 & 1 & 2 & 3 & 0 \end{array}\right\} = \frac{5}{3}$$

The array at the bottom is extended with a 0

$$y(1) = \left\{\begin{array}{ccccccccc} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & & & & & & \\ \times & \times & \times & & & & & & \\ 1 & 2 & 3 & 1 & 2 & 3 & 1 & 2 & 3 \end{array}\right\} = 2$$

$$y(2) = \left\{\begin{array}{ccccccccc} & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & & & & & \\ & \times & \times & \times & & & & & \\ 1 & 2 & 3 & 1 & 2 & 3 & 1 & 2 & 3 \end{array}\right\} = 2$$

$$\vdots$$
$$\vdots$$

$$y(7) = \left\{\begin{array}{ccccccccc} & & & & & & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ & & & & & & \times & \times & \times \\ 1 & 2 & 3 & 1 & 2 & 3 & 1 & 2 & 3 \end{array}\right\} = 2$$

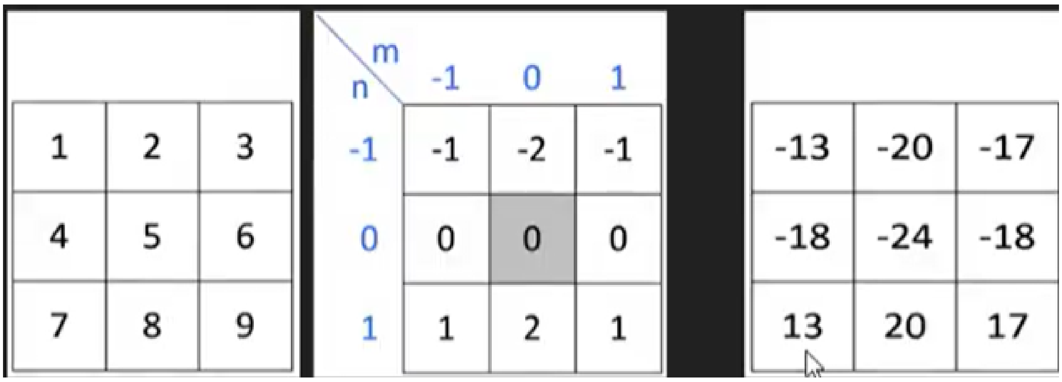
- The upper one is moving two steps

Two choice parameters:

1. Stride: How many steps the filter is moving
2. Padding: Putting extra 0s at the end and beginning

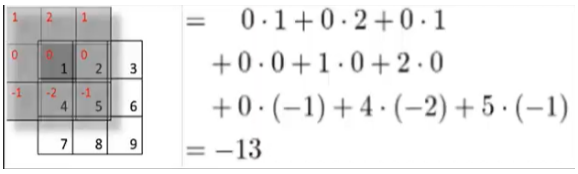
Convolution in 2D

- For an image, there are two dimensions. Height and width.

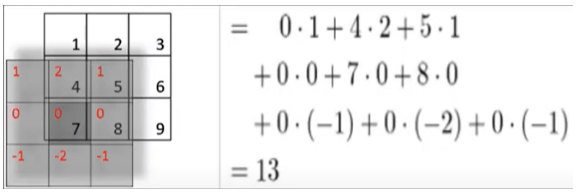


Images with 9 pixels

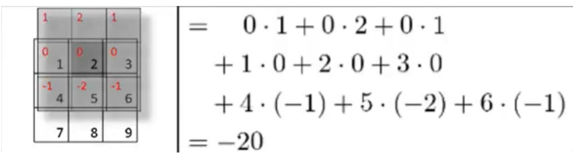
- The filter is called the **kernel.** It's convolved with the input image and the output is generated.
- We should decide the shape of the kernel. Usually square shape, but other shapes can be used.
- The kernel average is at the centre.



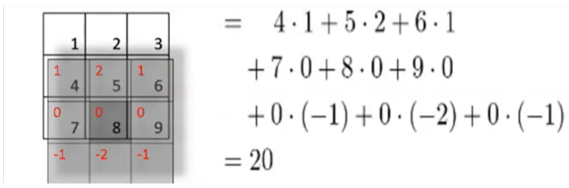
1



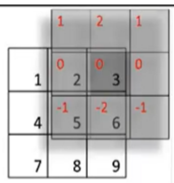
7



2

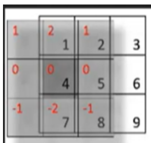


8



$$= 0\cdot1+0\cdot2+0\cdot1 +2\cdot0+3\cdot0+0\cdot0 +5\cdot(-1)+6\cdot(-2)+0\cdot(-1) = -17$$

3

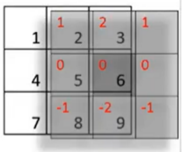


$$= 0\cdot1+1\cdot2+2\cdot1 +0\cdot0+4\cdot0+5\cdot0 +0\cdot(-1)+7\cdot(-2)+8\cdot(-1) = -18$$

4

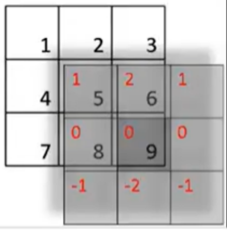
$$= 1\cdot1+2\cdot2+3\cdot1 +4\cdot0+5\cdot0+6\cdot0 +7\cdot(-1)+8\cdot(-2)+9\cdot(-1) = -24$$

5



$$= 2\cdot1+3\cdot2+0\cdot1 +5\cdot0+6\cdot0+0\cdot0 +8\cdot(-1)+9\cdot(-2)+0\cdot(-1) = -18$$

6



$$= 5\cdot1+6\cdot2+0\cdot1 +8\cdot0+9\cdot0+0\cdot0 +0\cdot(-1)+0\cdot(-2)+0\cdot(-1) = 17$$

9

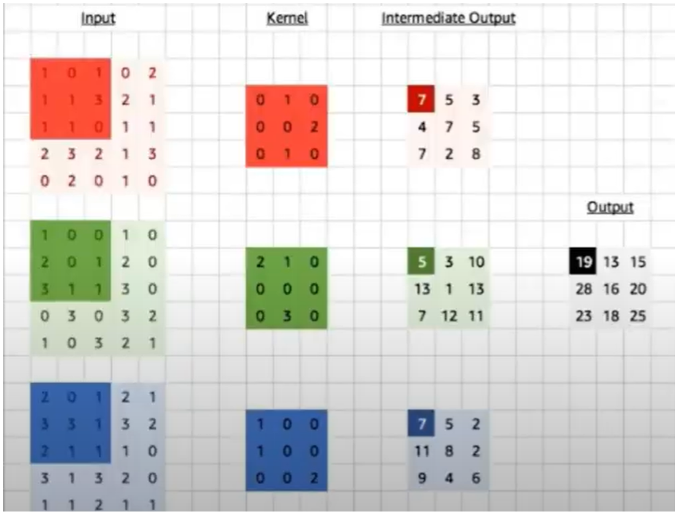
-13	-20	-17
-18	-24	-18
13	20	17

Result

Convolution in 3D

- Usually, we work with 3D colour images(RGB).
- Greyscale images are 2D and only have one channel.
- For colour images, we have 3 channels. Apart from height and width, we will need the depth/channel information.

How convolution is done



3D Convolution

- The channels are separated and then the filter/kernel is convolved exactly like 2D convolution.
- Afterwards, the channels are merged to get the final image.