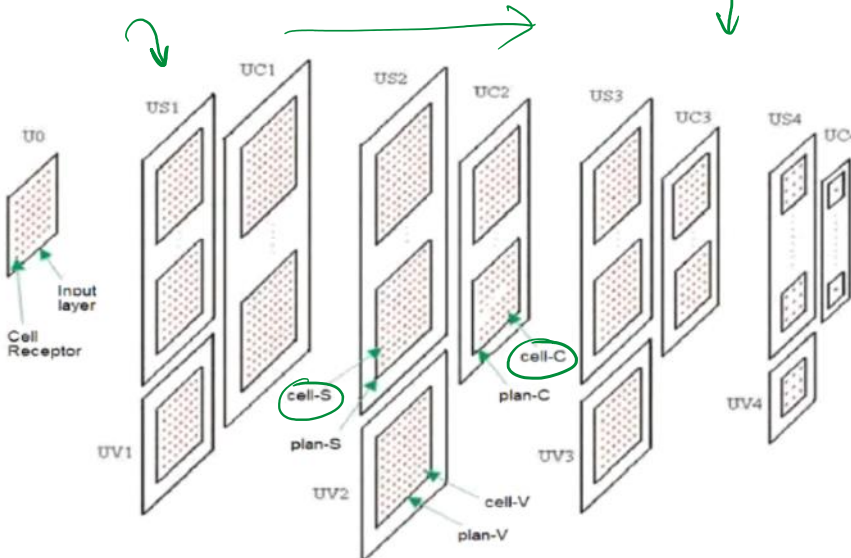Neocognitron → Fukushima

Yann Lecun → CNN → Backprop Convolution
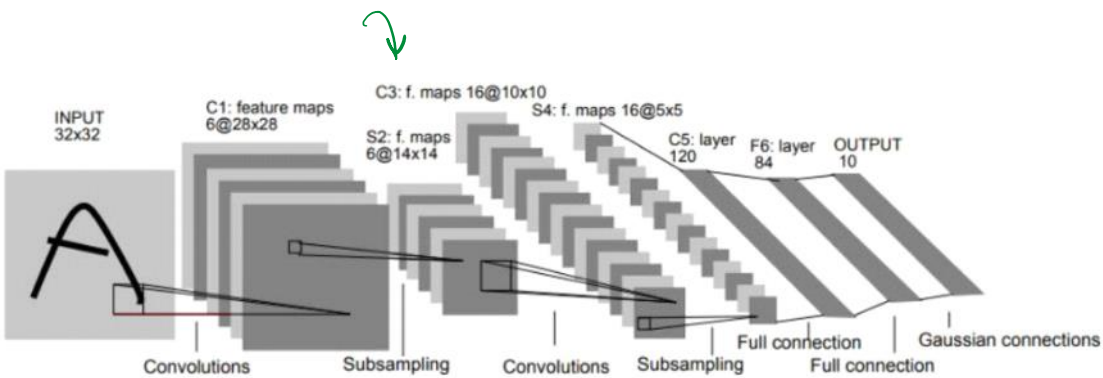
early CNN

2012 → AlexNET → ImageNET → CNNs
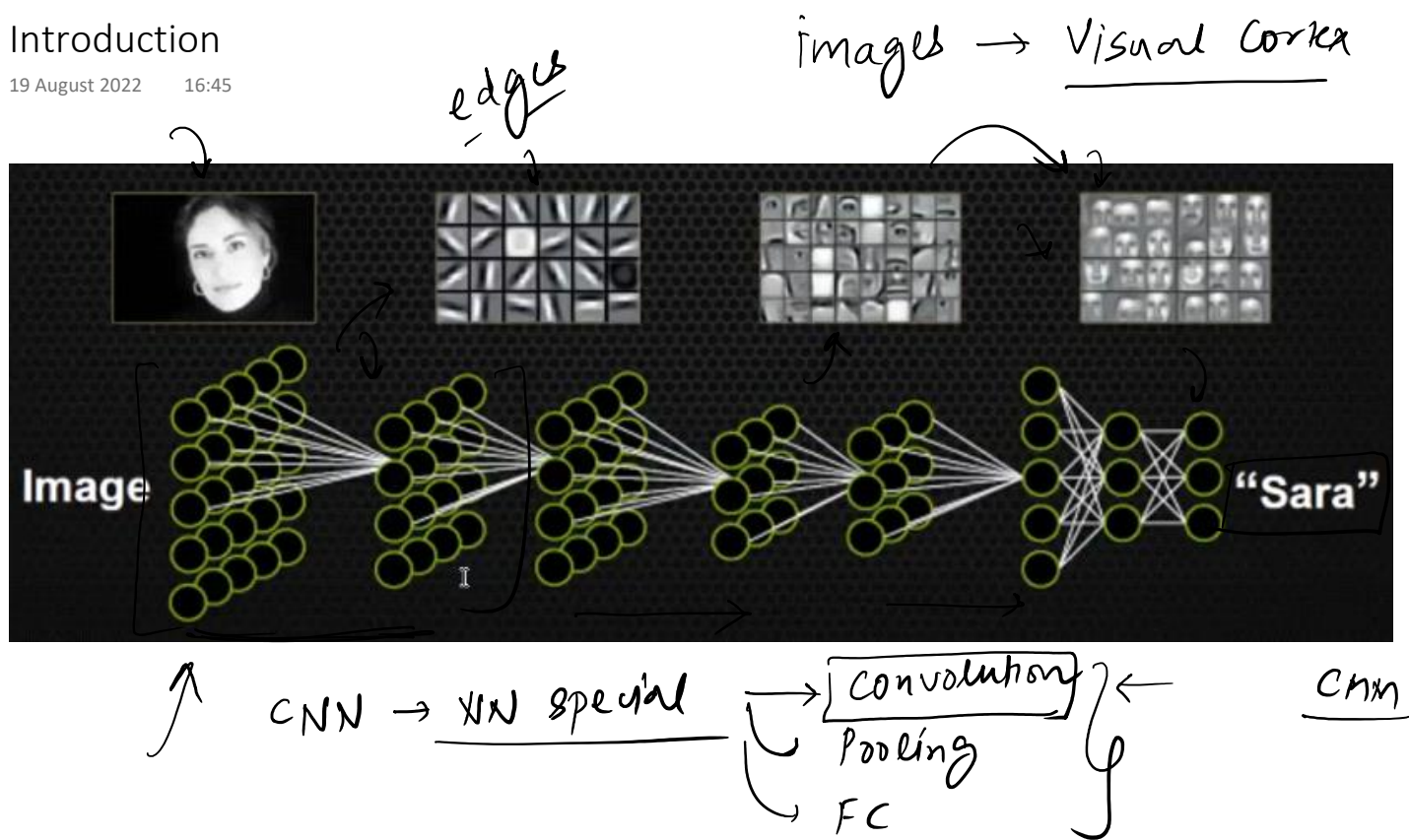
edge

images → Visual Cortex

Image → "Sara"

CNN → NN special → Convolution
                     Pooling
                     FC

CNN

Basics of Images

$0-255 \rightarrow \boxed{0-1}$

B/W
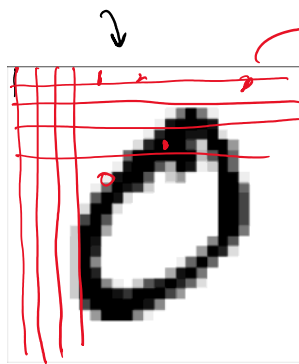
Greyscale → B/W   1 channel

RGB → colored imgs
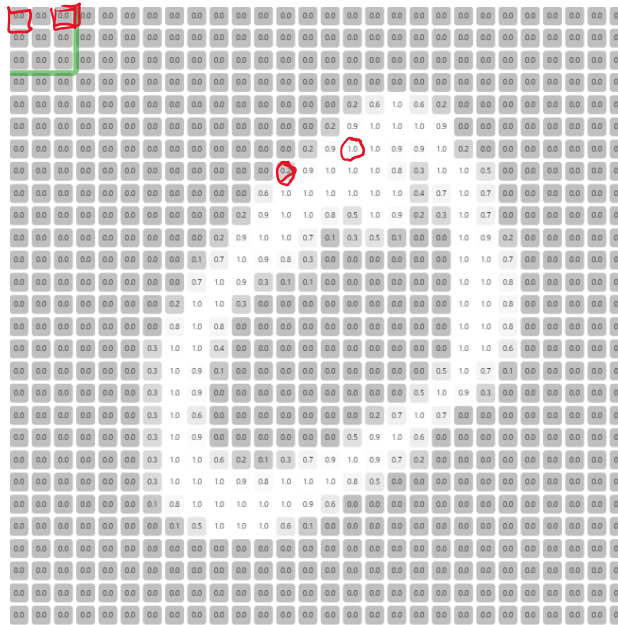
28 × 28 pixels

2D arrays →

2D numpy arrays

mnist → 0

(28 × 28)

3 channels → Red Green Blue (primary)
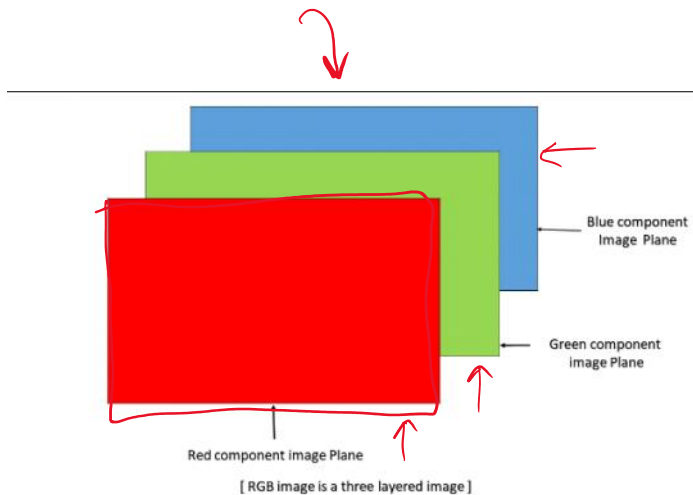
Blue component Image Plane

Green component image Plane

Red component image Plane

[RGB image is a three layered image]

$\boxed{228 \times 228 \times 3}$

3 channels

$255 \times 1 + 255 \times 1 + \ldots$ horizontal

horizontal film

edge detector

$-1 \times 0 + -1 \times 0 + -1 \times 0$

image

0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
255	255	255	255	255	255
255	255	255	255	255	255
255	255	255	255	255	255

↑ 6×6

*

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

=

ANN weights

filter/kernel

Matrix 3×3

feature map

0	0	0	0
255	255	255	255
255	255	255	255
0	0	0	0

0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
255	255	255	255	255	255
255	255	255	255	255	255
255	255	255	255	255	255

(6 × 6)

*

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

(3×3)

=

(4×4)

$(28 \times 28) \longrightarrow (26 \times 26)$
$n \times n$
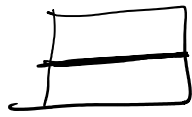
$(3 \times 3) \longrightarrow$? (26×26)
$m \times m \longrightarrow (n - m + 1) \times (n - m + 1)$

$(64 \times 64) \quad (3 \times 3) \longrightarrow (62 \times 62)$

Demo

Working with RGB Images

$3 \times 3 \times 3$ 3 channel

$3 \times 3 \times 3 = 27$ values

6 x 6 x 3   *   **3 x 3 x 3**   =   **4 x 4** 4×4 single channel

27   27

feature map
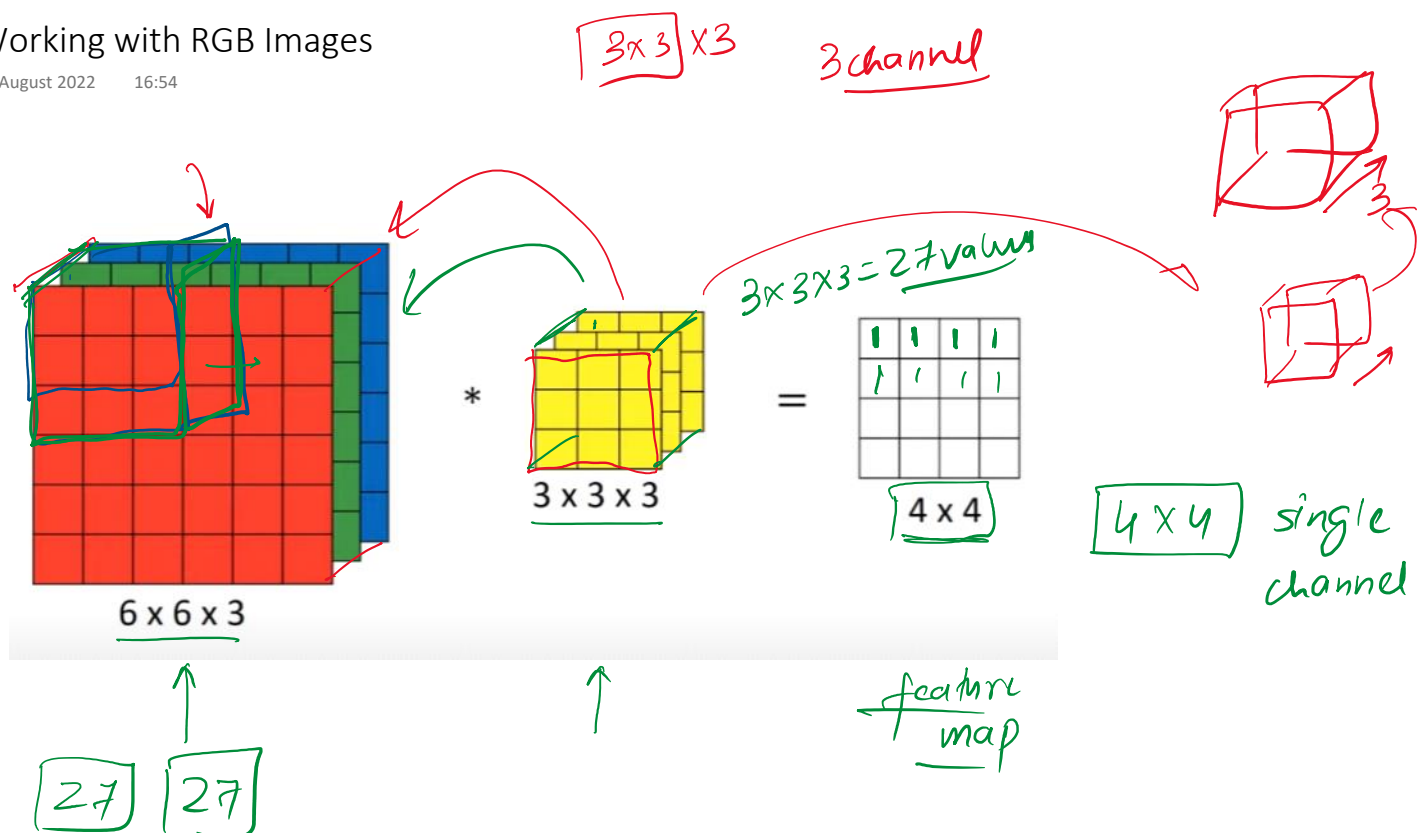
$$\boxed{m \times m \times C} \quad \boxed{n \times n \times C} \longrightarrow \frac{(m - n + 1)\;(m - n + 1)}{\text{single channel}}$$

Multiple Filters

23 August 2022   08:24

Vertical edge

$3 \times 3 \times 3$

Horizontal edge

$3 \times 3 \times 3$

=

4×4
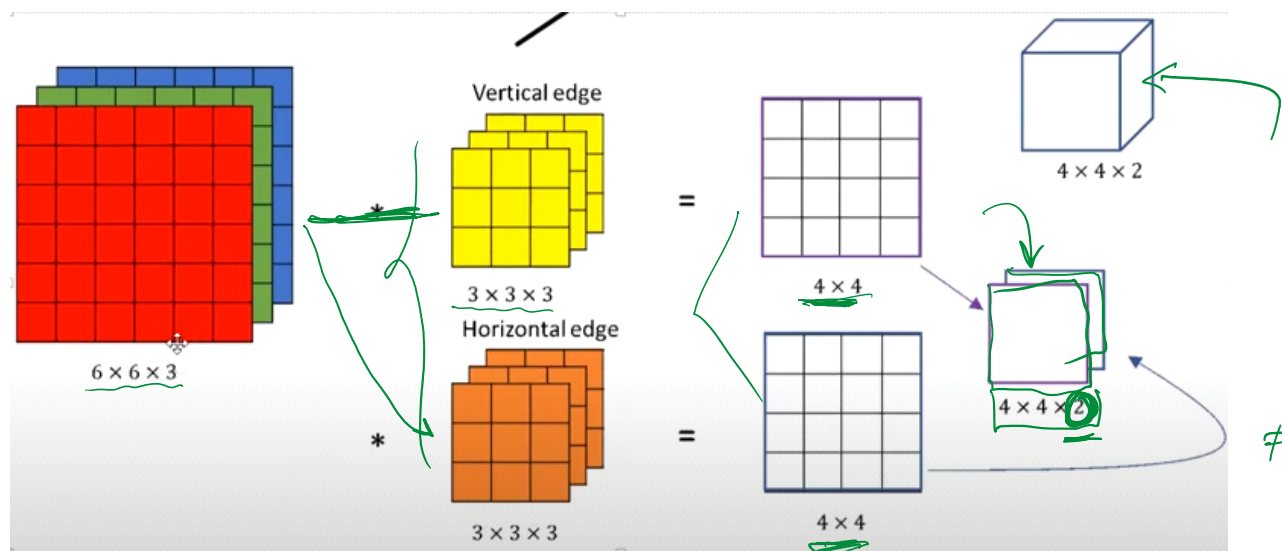
=

4×4

$4 \times 4 \times 2$

$4 \times 4 \times 2$

filters used

Image taken from Andrew NG's lecture

$6 \times 6 \times 3$

$10 \longrightarrow$

$4 \times 4 \times 10$