YOLOV8

1.5 ~ 2 months

2012 → 2023

DL Era

Check

Img Classification

X { LeNet, AlexNet, VGG Net }

lag or no lag

Agenda

| Resnet |

1) Computer Vision

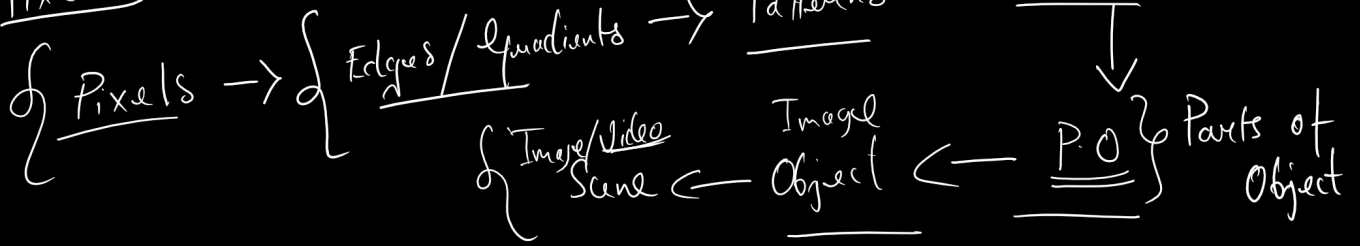VGG16

1) Kernels
2) Channels

How to create a custom kernel?

a

Filters | Kernel | Feature Extractors

$$\boxed{a\ B\ C\ d\ e\ D\ f\ F}$$

2 3 8

$\underline{\text{Pixels}} \rightarrow$ Smallest Component

$\{ \underline{\text{Pixels}} \rightarrow \{ \underline{\text{Edges}} / \underline{\text{Gradients}} \rightarrow \underline{\text{Patterns}} \longrightarrow \underline{\text{Textures}}$

$\{ \underset{\text{Scene}}{\text{Image/Video}} \longleftarrow \underset{\text{Object}}{\text{Image}} \longleftarrow \underline{\text{P.O}} \} \underset{\text{Object}}{\text{Parts of}}$

$\underline{\text{Video}} \qquad \underline{\underline{\text{Images}}}$

$\overset{\curvearrowleft}{\underline{h \times w \times 3}}$

⑤ steps $\longrightarrow$ ⑤ $\underline{\text{Layers}}$

$\underline{\text{Image RGB}} = \underline{h \times w \times 3}$ $\downarrow$

$1 \times 1 \times \underline{64}$
$\overline{216}$
$\overline{512}$

$\underline{\text{Small} \qquad \text{to} \qquad \text{Big}}$

$\underline{\text{Pixel} \qquad \text{to} \qquad \text{Scene/Object}}$

$\underset{\text{2012}}{\underline{}}$

$\underline{\text{Big}} \qquad \underline{\text{to}} \qquad \underline{\text{Small}}$

$\underset{2008}{} \quad \underset{2012}{} \quad \underline{DL}$
$\overline{2017}$

$\text{Image} \longrightarrow \underline{\text{Object}} \twoheadrightarrow \underline{\text{Parts of Object}} \longrightarrow \underline{\text{Textures}} \longrightarrow \text{Pattern}$

$\downarrow$

Edges
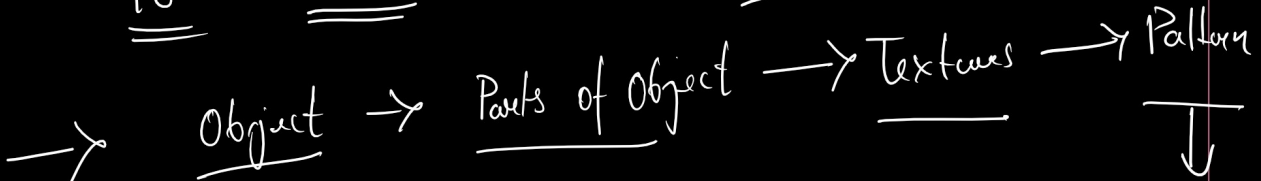
$\underline{\text{Convolutional}}$

$200 \times 200 \times \boxed{12}$

$\downarrow$

Pixels

$200 \times 200 \times ③$

$200 \times 200 \times 12$

channels/ feature Maps

12 channels

Image ↘
X Colour Features

Feature Maps ↘
No Colour ↘
Features Maps

Filters | Kernel | Feature Extractions

Input

200 X 200 X 12

12 filters

= output Channel

No of Kernel

Kernel

3 X 3

h x w

$\{$ 11 X 11

7 X 7

3 X 3

& 5 X 5 1 X 1

Complexity
of Features

3 X 3

= less params

3 X 3

= 9

5 X 5

= 25

Input
36 X 36 X ②

5 X 5 X ⑫

Convolution   Input
36 X 36   X ②

X   12

(Conv) ① 3 X 3

Output = 32 X 32 X 12

3×3
$= \underline{\underline{9}}$

$9 + 9 = \underline{\underline{18}}$

$\underline{\text{Input}} = \dfrac{34 \times 34 \times ?}{(\text{Conv}) \ ② \ 3 \times 3 \ \times ?}$

$= \ 32 \times 32$

$5 * 5$
$= 25$

Ⓘ $360 \times 360 \times ?$

Ⓒ $^{\text{Conv}} \ \boxed{3 \times 3} \times ? \ = 9$

Ⓞ $\ 358 \times 358 \times ?$

Single Conv Operation

$358 \times 358 \times ?$

$3 \times 3 \ \times ? \ = 9$

$\boxed{356 \times 356} \times ?$

Stride
① ②

3×3

$3 \underbrace{(3 \times 3)}$
$\quad = 9 \times 3$
$\quad = 27$

Ⓣ $360 \times 360 \times ?$

Ⓒ $5 \times 5 \times ?$

Ⓞ $\boxed{356 \times 356} \times ?$

$3 \times 3 \times 2$
$= ⑱$

5×5
$= \underline{\underline{25}}$

$\underset{358}{\overset{360}{\downarrow}} \quad \underset{356}{\overset{360}{\downarrow}}$

Information
loss

7×7

7×7
$= 49$

<u>cooked up</u> values

<u>black pixels</u>

<u>ANW</u>
Ly
$?$

3

Input _____ $h \times w \times \boxed{3}$

Output $1 \times 1 \times 256$

                Before Flatten

360 × 360

100 × 100

50 × 50

whole
whole
whole

Image
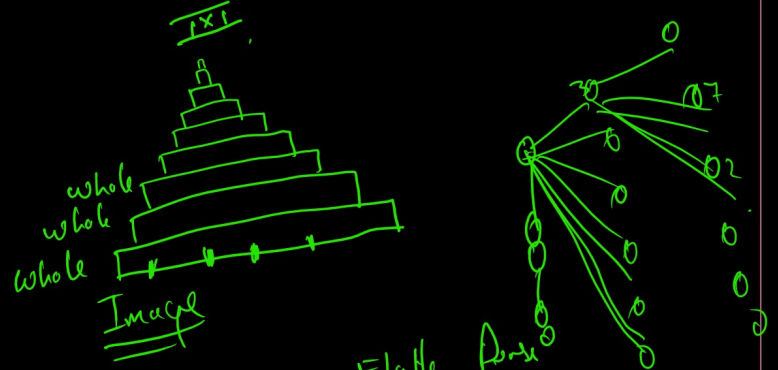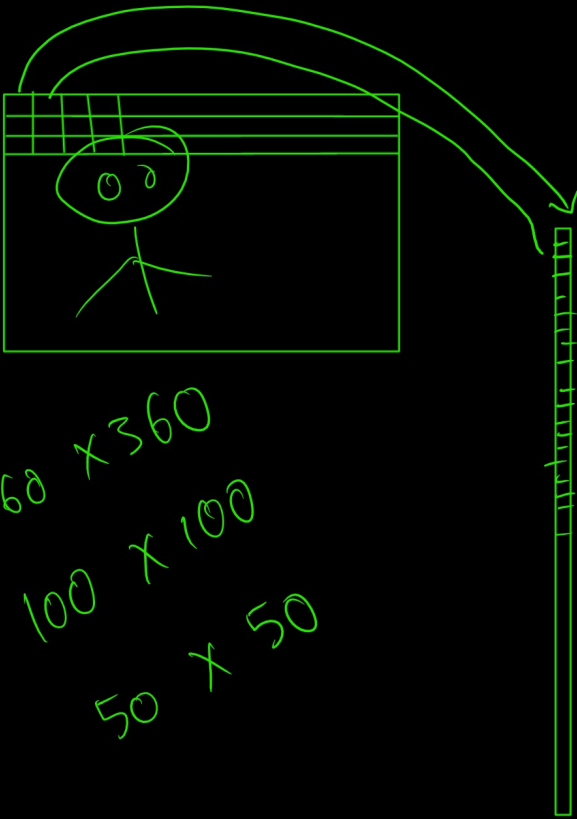
Dense Dens Flatte Dense

↑
Arrange the pixels

Vertically

Con
Max

Dense Dense Flatten Dense

ANN

Fully

Input

Dense
Dense
Dropout
Dense

Limt
Abount
Value

Flatten   Dense (10)
                Softmax

Example:

Convention }   Starting   256×256× ?
  Input   96×96
  C
  C Max Pooling
  C
  C Max Pooling   Reach as close
  F   Closing   (1×1)
X FC X
 Softmax }   (7)

4×4   X 256
  =
     (1×1)   2×2×25
              2×2

Max
 C
Flatten
 FC
Softmax

4×4
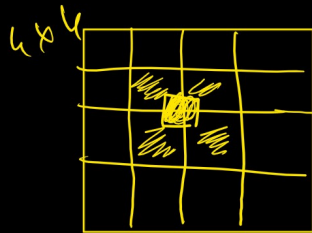
= 16

= 9
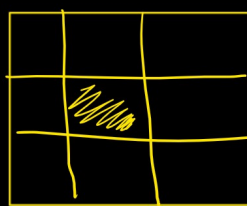
Odd size filter

X (Even Size filter)

Symmetry Issues   36 × 36 × 3

X (29) X (29)   = (8×8)   3×3

  (24 × 24) → Mathematically
                iF is fine

Con   CNNs   Alexnet   leNet   Vccu   Resnet

②

{ Parameter Calcution }

{ CNN Ruvision }

PW

Image Classification

2 : ①