(1) rapid advances in CV are enabling brand new applis to view, (few yrs ago ↓)

(*) invent new products & apply.

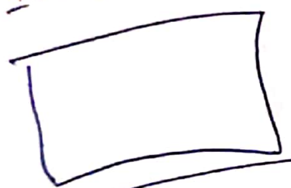
(2) (5) even if u don't end up building CV systems
CV research imp
with NN architectures & algos

CV ← Speech recognition
cross fertilization

CV ⇒ Speech recognition
ideas →

CV problems

Img



Challenges, i/p big imgs ⇒ 12288 i/p features has dimention

(1) cat img 64x64 img x3 (RGB color channels)
Small img.

1 mega pixel.

(2) 1000 X 1000 X 3 ⇒ 3 million.

CV ⇒ tiny imgs
Big imgs ⇒ CNN.

CNN

ex:- edge detection ⟹ specify how to find vertical edges in an img.

* convolution operation is one of the BBB of CNN

* detect edges?

① detect vertical edges in the img.

② " horizontal edge — —

ex:- Vertical edge detection ⟹ a 3x3 region since we are using 3x3 filter bright pixels on left.

6×6 × 1 (Grey)

construct 3x 3 matrix (filter) ⟺ (Kernel)

* Symbol for convolution

convolution o/p 4x4 img.

ele wise product.

1	0	-1
1	0	-1
1	0	-1
bright pixels		dark pixels.

python :- conv-forward

Video1 : CV, imgs, tiny CV ← large CNN

Video2 :- convolution operation (kernel or filter) edge detection

(V3) (2)

=) diff b/w +ve & -ve edges
light & dark

=) types of edge detectors =) algo learn

-ve :-) dark to light rather than a light to dark transition.

Filter does make a diff b/w light to dark

horizontal :- bright top dark bottom.

45, 70, 75°

pad

(V3)

(V4)

vertical edge detection	horizontal edge detect.	45°, 70°, 75°

Padding:

* In order to build deep NN one modification to the basic convltn operation that u need to really use is "PADDING"

* 6×6 * 3×3 $\Rightarrow$ 4×4
  $n\times n$ $f\times f$ $\Rightarrow (n-f+1)\times(n-f+1)$
  e.g. 6×6 3×3 $\Rightarrow (6-3+1)=4\times4$

Pblms :- (1) After applying convltn $\Rightarrow$ everytime ing shrinky
(2) pixel at the corner or the edge, used only in one of the o/ps
middle pixel $\Rightarrow$ so many times.

$6\times6 \xrightarrow{\text{pad}} 8\times8 \xrightarrow{\text{convol } 3\times3} 6\times6$

p = padding = 1 o/p $(n+2p-f+1)\times(n+2p-f+1)$

* Valid & Same Convolutions

Valid :- (NO padding) $\underset{6\times6}{n\times n} * \underset{3\times3}{f\times f} \Rightarrow \underset{4\times4}{(n-f+1)\times(n-f+1)}$

Same : pad so that o/p size is same as i/p size.

$n\times n$ ing $\Rightarrow (n+2p-f+1)\times(n+2p-f+1)$

if $(n+2p-f+1)=n \Rightarrow \boxed{p=\frac{f-1}{2}}$

if f = odd (1) you can make sure that the o/p size is same as i/p. (2) f is odd always
p = even (0) $\Rightarrow$ Assymetric padding

ex: 5×5 $\Rightarrow$ central pos.

Strided Convolutions :-

ex:

$7 \times 7 * 3 \times 3 = 3 \times 3$

stride = 2 → jump 2 right positions.

$$\left\{ \begin{array}{l} n \times n * f \times f \\ \text{padding } p \quad \text{stride } S \end{array} \right\} \Rightarrow \left[\left(\frac{n+2p-f}{S} \right) + 1 \right] * \left[\left(\frac{n+2p-f}{S} \right) + 1 \right]$$

narrow operation } = cross correlation. fraction → round.

V6 ⇒ Strided Convolutions correlation.

Convolutions over volumes.

(V7:) Convolutions over volumes.

$6 \times 6 \times 3 * 3 \times 3 \times 3 = 4 \times 4 \times 1$

height / width / channel

stride by 1

One layer of a Conv n/w

add bias b_1 + non-linearity.

$6 \times 6 \times 3 \xrightarrow{a^0} 3 \times 3 \times 3 \Rightarrow 4 \times 4 \times 2 \xrightarrow{a^1}$ } one layer of Conv net

$4 \times 4 \times 10$ ⇒ 10 filters are (10)

Pblm ⇒ filter ⇒ $3 \times 3 \times 3$

27 parameters. + bias (28 parameters)

2 3 . . . 10 filters

$10 \times 28 = 280$ parameters

1 layer :-

$f^{[l]}$ = filter size

$p^{[l]}$ = padding

$s^{[l]}$ = stride

$n_c^{[l]}$ = no. of filters.

each filter $f^{[l]} \times f^{[l]} \times n_c^{[l-1]}$

I/P :- $n_H^{[l-1]} \times n_w^{[l-1]} \times n_c^{[l-1]}$ (no. of channels)

(Ht) (width)

$n_H^{[l-1]} \times n_w^{[l-1]} \times n_c^{(l-1)}$

O/P :- $n_H^{[l]} \times n_w^{[l]} \times n_c^{[l]}$ → size of O/P

$$n_{(H)_w}^{[l]} = \left\lfloor \frac{n_H^{[l-1]} + 2p^{[l]} - f^{[l]}}{s^{[l]}} + 1 \right\rfloor$$

weights

bias

$4 \times 4 \times 10 \Rightarrow$ Activation $= a^{[l]} \Rightarrow n_H^{[l]} \times n_w^{[l]} \times n_c^{1}$

$A^{[l]} \rightarrow m \times n_H^{[l]} \times n_w^{[l]} \times n_c^{[l]}$

(V5)

Strided Convolution :-

ex:

7×7 * 3×3 = 3×3

stride = 2 — jump 2 right position.

$$\left\{ \begin{array}{l} n \times n \quad * \quad f \times f \\ \text{padding } p \quad \text{stride } S \end{array} \right\} = \Rightarrow \left[\left(\frac{n + 2p - f}{S} \right) + 1 \right] \times \left[\left(\frac{n+2p-f}{S} \right) + 1 \right]$$

fractional → round.

narrow operation } = cross correlation.

V6 ⇒ Strided Convolution correlation.

V7: Convolution over volume.

6×6×3 * 3×3×3 = 4×4×1

height / width / channel

slide by 1

V9

A Simple Convolution n/w example- convnet

ing classification

cat or not;

$39 \times 39 \times 3$

$n_H^{[0]} = n_w^{[0]} = 39$

$n_c^{[0]} = 3$

$\Rightarrow$ $f^{[1]} = 3$, $s^{[1]} = 1$, $p^{[1]} = 0$, 10 filters.

$37 \times 37 \times 10$

$\frac{n+2p-f}{s} + 1 = \frac{39+0-3}{1} + 1 = 37$

$n_H^{[1]} = n_w^{[1]} = 37$

$n_c^{[1]} = 10$

filter = 5

$f^{[2]} = 5$

$s^{[2]} = 2$

$p^{[2]} = 0$

filters = 20

$17 \times 17 \times 20$

$\left[\frac{(n+2p-f)}{s} + 1\right] = 17$

$\left[\frac{37 + (2*0) - 20}{2} + 1\right] =$

$n_H^{[2]} = n_w^{[2]} = 17$

$n_c^{[2]} = 8$

$\Rightarrow$ $f^{[3]} = 5$, $s^{[2]} = 2$, filters = 40, $p^{[2]} = 0$

$7 \times 7 \times 40$

1960

1960

y

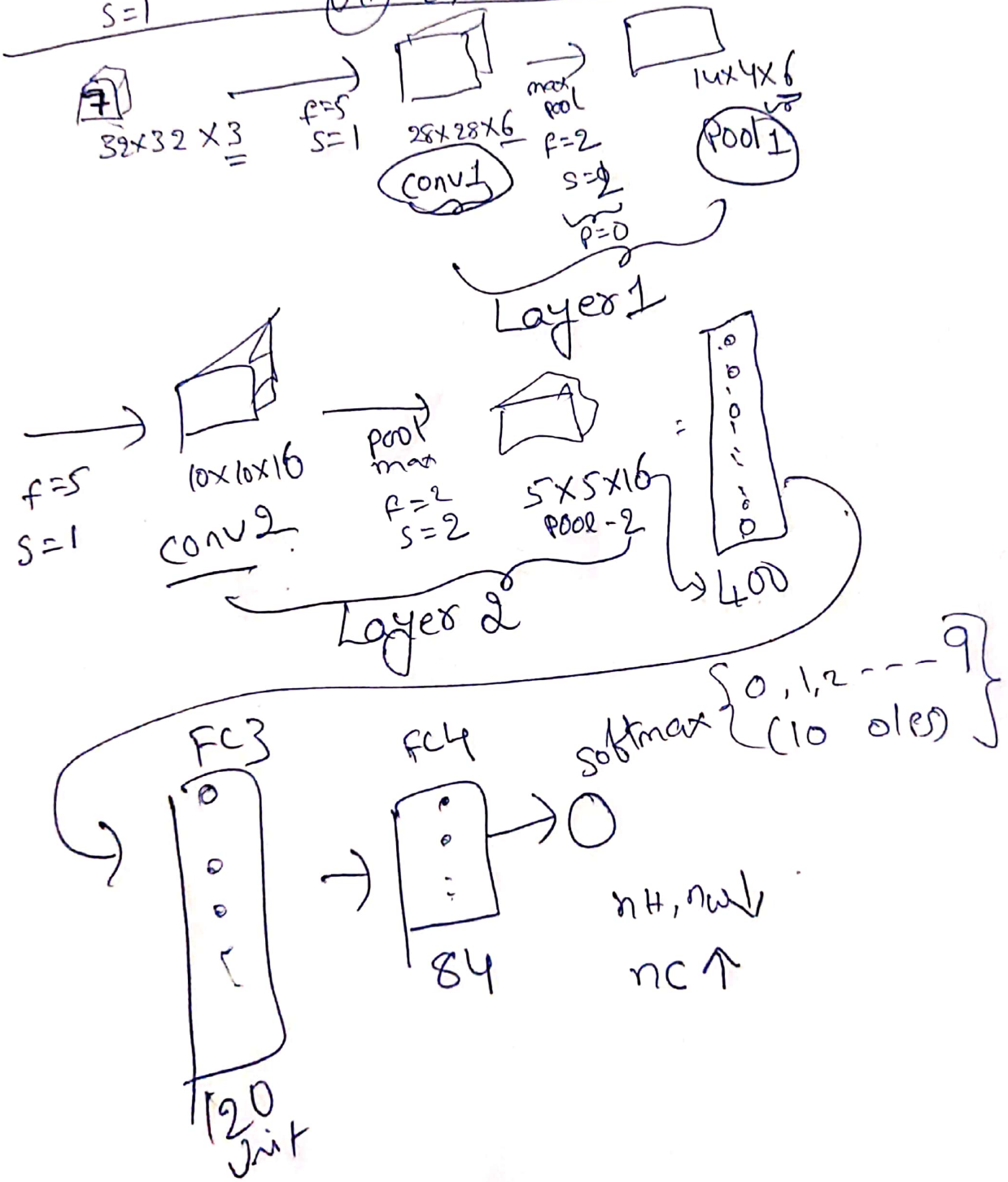
logistic / Softmax

(v10) Pooling Layers → no parameters
→ fixed function

max pool:-
4×4 → 2×2
5×5 ⇒
f=3
s=1

avg pool
7×7×1000 ⇒ 1×1×1000

(v11) Example. (Lenet-5)



Why convolution

1. Parameter sharing
2. Sparsity of connection

Quiz

2. 300 × 300 × 3

100 neurons

1
0
!
100

+1

63 × 63 × 16

32 filter 7 × 7

s = 2

p = 0

n = 63

s = 2

p = 0

f = 5

$\frac{63 - 25}{2}$

n

15 × 15 × 8

p = 2

63 × 63 × 16 f = 32

7 × 7

s = 1

32 × 32 × 16

s = 2

f = 2

16 × 16 × 16