

CNN :- Convolutional Neural Network.

Key concept :-

① Input (Image) \rightarrow Pooling \rightarrow Flattening \rightarrow Output
(It is downsampling tech.)

Pooling - It is downsampling tech use in CNN to reduce the size of image without losing imp. feature.
 \Rightarrow computational fast & reduce overfitting.

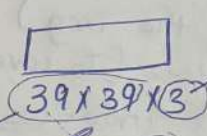
Flattening :- Convert image \rightarrow 1D (easy for classification)

Before fully connected layer.

It process images into matrices

\Rightarrow Neural N/W like fully connected layer no input dense like imp. layer.
 \Rightarrow model use imp. feature & give final decision.

$$\text{output size} = \frac{n+2p-f}{s} + 1$$

♀  channel filter = always odd
image size $39 \times 39 \times 3$ filter size 3×3
Stride = 1, Padding = 0, No. of filter = 10, $n = 39$, filter = 3

Step 1 :- $\text{output size} = \frac{n+2p-f}{s} + 1 \Rightarrow \frac{39+2 \times 0 - 3}{1} + 1$
 $\Rightarrow 39 - 3 + 1 \Rightarrow 37$

$$= 37 \times 37 \times 10$$

Step 2 :- For deep extracting & extracting imp. feature up will increase filter size, stride

filter size = 5×5 , stride = 2, $n = 37$, filter $f = 5$

$$\text{output} = \frac{n+2p-f}{s} + 1 \Rightarrow \frac{37+2 \times 0 - 5}{2} + 1$$

$$= 16 + 1 = 17 \Rightarrow 17 \times 17 \times 20$$

Step 3 :- ~~some~~ some surif padding increas krung.

$P = \frac{f-1}{2}$ even no.
 $S = 2$ filter 5×5 , $n = 17$, $f = 5$

output = $\frac{n + 2P - f + 1}{S}$
 $= \frac{17 + 2 \times 2 - 5}{2} + 1 = \frac{17 + 4 - 5}{2} + 1 \Rightarrow \frac{16}{2} + 1$

$\Rightarrow 9 \times 9 \times 40$

{ \therefore step 1 se step 3 tak pooling hua }

Step 4 :- calculate flatten

Given 120 neuron -

$9 \times 9 \times 40 \times 120 + 120$ ^{neuron} ^{Bias}

imp feature = 388920

CNN :- ① understanding the image ism pura image ko cover krung output size niklung.

- ② Pooling \rightarrow ① output size niklung (To understand the img)
 ② filter size & stride increas krung (To cover imp feature)
 ③ padding increas krung (To avoid info-loss)

③ calculate flatten

Max Pooling :- ~~store~~ down sampling tech. Cover only most imp. feature of image.

Calculate one time output size =

$32 \times 32 \times 3$, $S = 2$, $f = 2$, $n = 32$, $P = 0$

output = $\frac{n + 2P - f + 1}{S}$ $\Rightarrow \frac{32 + 2 \times 0 - 2}{2} + 1 \Rightarrow \frac{32 - 2}{2} + 1$

$\Rightarrow 15 \times 15 \times 3$

$15 \times 15 \times 3$

$\Rightarrow 15 + 1 = 16$

LeNet-5 Arch - CNN + max pooling.

- # CNN (extract feature)
- # max pooling (reduce size)
- # fully connected (final decision)
- # softmax layer (which digit)

Q $32 \times 32 \times 3$ $f = 5 \times 5$, $f = 5$, $s = 0.1$, $P = 0$, $n = 32$
 f no. of $f = 8$

CNN output size = $\frac{n+2P-f}{s} + 1 = \frac{32+2 \times 0 - 5}{1} + 1$
 $= \frac{27}{1} + 1 \Rightarrow 28$
 $28 \times 28 \times 8$

max pooling = $\frac{n+2P-f}{s} + 1 \Rightarrow \frac{28+2 \times 0 - 2}{2} + 1 \Rightarrow \frac{26}{2} + 1$
 $= 14$
 $14 \times 14 \times 8$ $f = 5$, $s = 1$

CNN:- $\frac{n+2P-f}{s} + 1 \Rightarrow \frac{14+2 \times 0 - 5}{1} + 1 \Rightarrow 9 + 1$
 $\Rightarrow 10$

Q $10 \times 10 \times 16$

max padding = $s = 2$, $f = 2$, $P = 0$
 $= \frac{n+2P-f}{s} + 1 \Rightarrow \frac{10+2 \times 0 - 2}{2} + 1 \Rightarrow \frac{8}{2} + 1$
 $= 5$

$5 \times 5 \times 16$

Fallting = 120, 84 fully connected layers

120:-

$5 \times 5 \times 16 \times 120 + 120 = 48120$

84:-

$84 \times 120 + 84 = 10164$

final = $1 \times 84 + 1 = 85$

Same Process for CNN ke layer

~~$32 \times 32 \times 3$~~ , ~~$28 \times 28 \times 8$~~ , ~~$10 \times 10 \times 16$~~

CNN layer Flatten:

$$5 \times 5 \times 8$$

$$\times 3$$
$$\rightarrow (5 \times 5 \times 8) \times 3 =$$

$$(5 \times 5 \times 16) \times 3 =$$

$$\rightarrow 5 \times 5 \times 16 \times 3 =$$

$$5 \times 5 \times 8 =$$

$$5 \times 5 \times 3 \times 8 \times 3 =$$

$$5 \times 5 \times 16 =$$

$$5 \times 5 \times 3 \times 16 \times 3 =$$

Total =