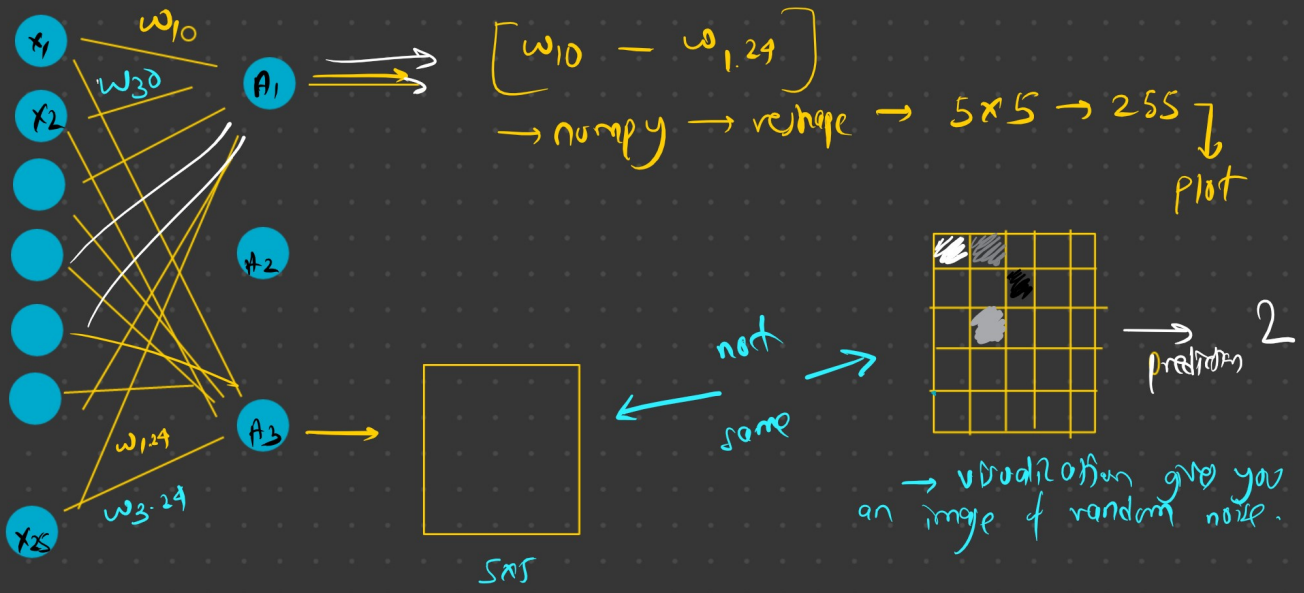
Agenda – 7 Dec 2024

Previous → ANN

Today's → CNN – Convolution Neural Network

→ flattening → ANN



ANN – Disadvantage

→ ANN is not focusing on features of image but finds features which are responsible for prediction.

→ Image – 720×720 → RGB
– ANN → 1,555,200 → 720 × 720 × 3
→ 15 lakh 55 thousand 200 neuron

→ Higher Computing is required.
→ Not optimized for images.

CNN → are a type of deep learning model which is effective for image processing & other task that are based on grid-like data.

Components of CNN

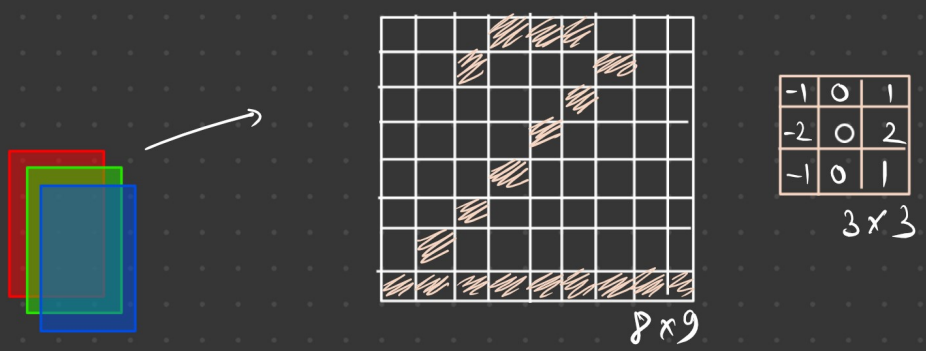
→ Convolution layer

→ Pooling layer

→ Fully connected layer [FC layer]

(1) Convolution layer

– core building blocks of a CNN. It is responsible for detecting local features in the input data such as edges, shapes & textures.



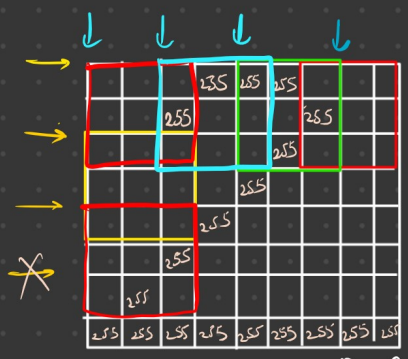
→ Convolve function

→ Stride

→ Kernel → matrix of weights

→ Sobel filter

→ Horizontal edges

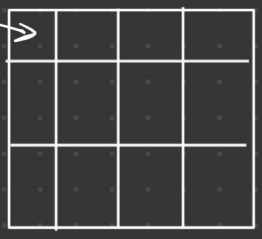


8x9

-1	0	1
-2	0	2
-1	0	1

3x3

$-1\times0+0+255$
$-510+0+0$
$0+0+0$



3x4

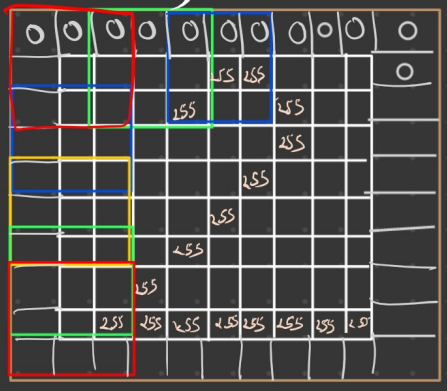
→ important features based on your kernel

4k → 4k
→ HD → only imp feat
→ ↓ computation

→ processed → 7
Actual → 2

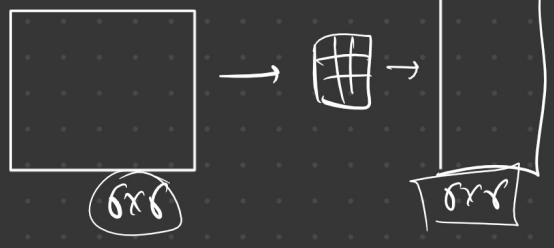
→ reduce the stride → 1
→ Advantage → cover more of image
→ with stride 2, we will lose information

→ Padding



→ we might miss features at the corners of the image
→ size of kernel
→ stride
→ stride = 2

→ Valid padding (No padding)
→ Same padding



padding = 0 → valid

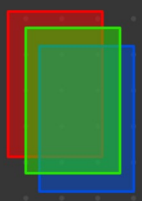
padding = "same" → same

= 1 → padd 1 pix with 0

= 2 —"— 2 —"—

= 3 —"— 3 —"—

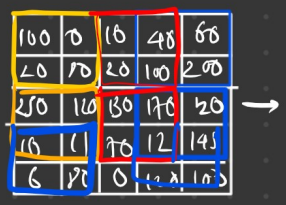
Pooling



→ R → Convolution →

→ stride = 2

→ kernel = 3x3



5x5

→

2x2

kernel

stride = 2

max pool

1.5 million → 1.2 million → pooling

input conv

→ reduce dimension

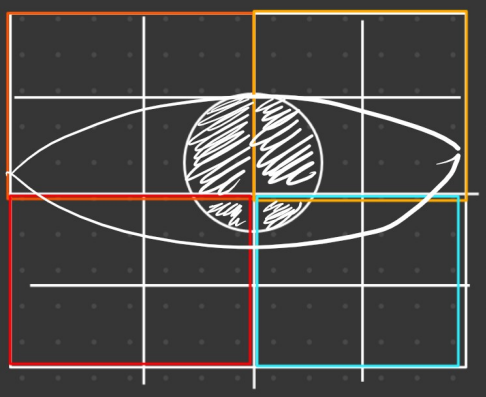
→ prominent feature

100 100 200

val val val

val val val

3x3



Different kinds of pooling

→ Max pooling

→ Min pooling

→ Average pooling

Input image → 60 x 60

— kernel → 3 x 3

— padding → 1

— stride → 2

→ pooling → 2

Image → Conv → op 1 → max pooling → op2

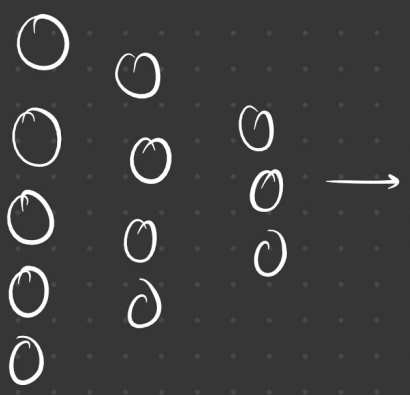
$$\text{output} \rightarrow \frac{\text{Input size} - \text{kernel/pool size} + 2 \times \text{padding}}{\text{stride}} + 1$$

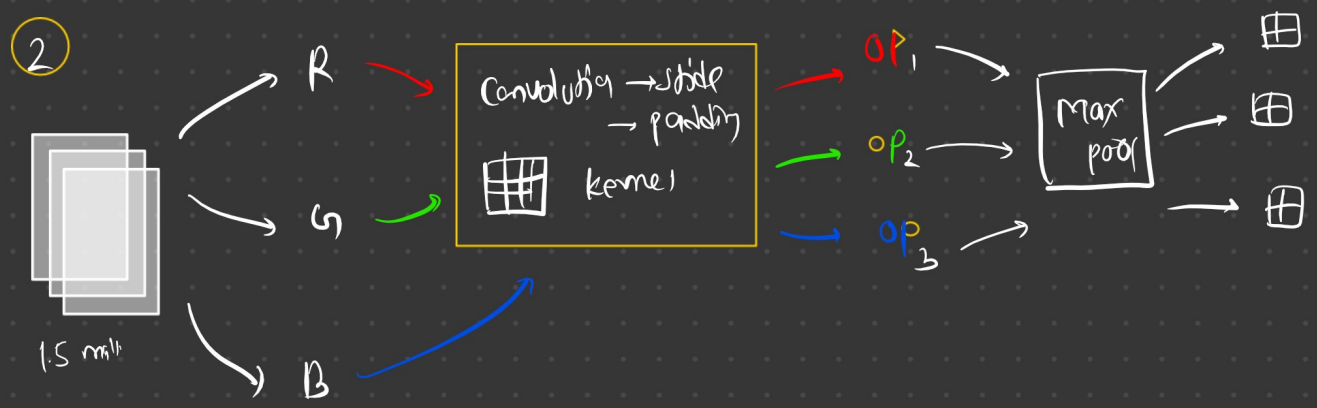
$$\text{op1} \rightarrow \frac{60 - 3 + 2 \times 1}{2} + 1 \rightarrow \boxed{30}$$

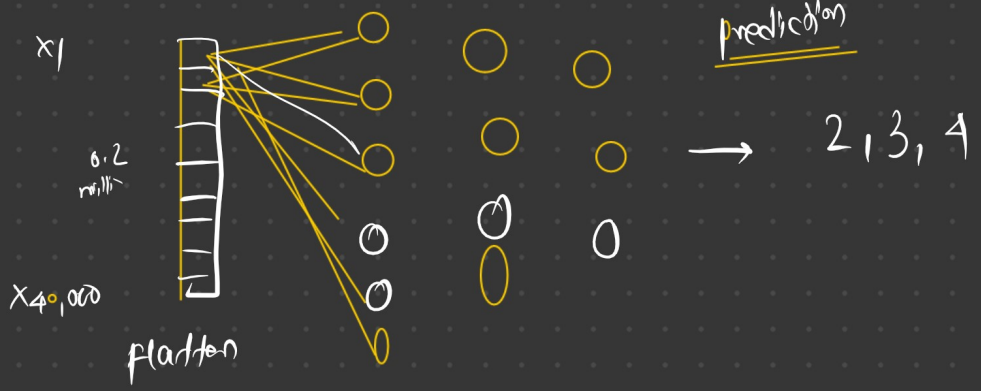
$$\text{op2} \rightarrow \frac{30 - 2 + 2 \times 0}{2} + 1 \rightarrow \frac{28}{2} + 1 \rightarrow \boxed{15}$$

FCL

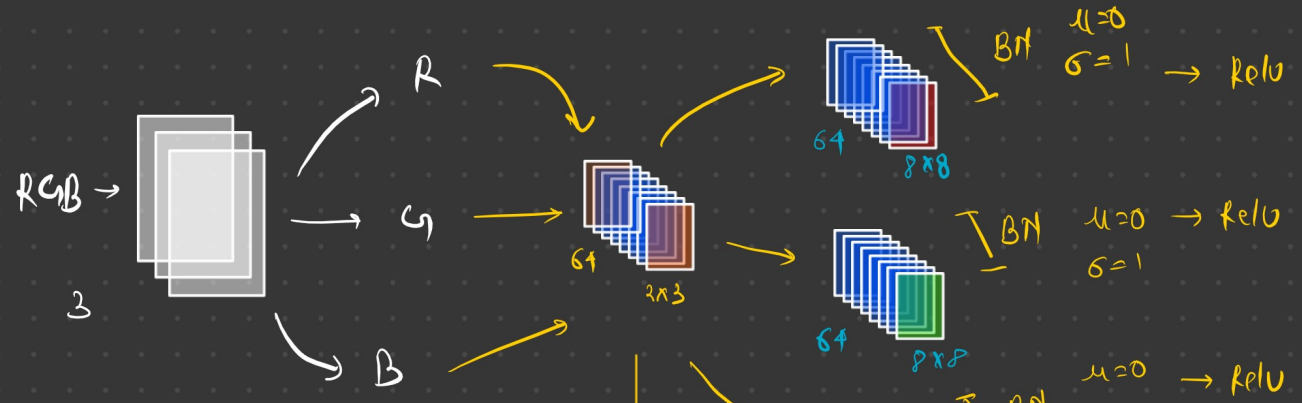
Fully connected layer







→ Extract relevant feature or prominent feature
→ To reduce dimension of the image



→ Separate channel
→ initialize lot of kernel.
→ 3x3
→ 64
→ Filter → 3x3x64

initialization random