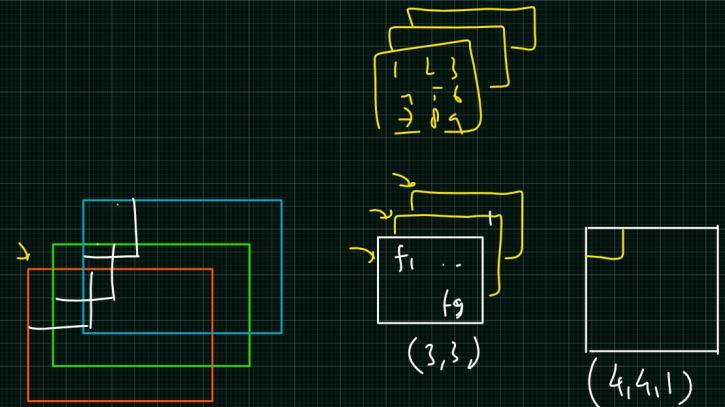
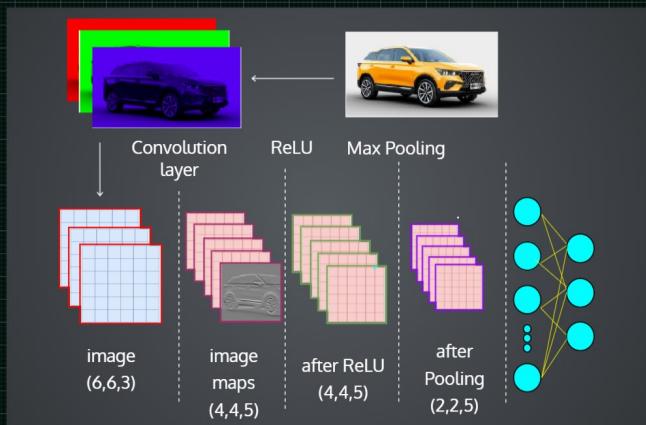


# AGENDA

- 1> CNN (Theory / Practical)
- 2> Full CNN implementation



$$\begin{aligned}
 Z_1 = & f_1 r_1 + f_2 r_2 + \dots \} \text{ red} \\
 & + f_1 g_1 + f_2 g_2 + \dots \} \text{ green} \\
 & + f_1 b_1 + f_2 b_2 + \dots \} \text{ blue}
 \end{aligned}$$

$S=1$

$$Z_L = f_1 r_1 + \dots$$

$$z_L = f_1 r_3$$

## Summary of Conv Layer

$$\text{input image size} = W_1 \times H_1 \times D_1$$

$$\text{output image size} = W_2 \times H_2 \times D_2$$

$$W_2 = \frac{(W_1 - F + 2P)}{S} + 1$$

$$H_2 = \frac{(H_1 - F + 2P)}{S} + 1$$

$$D_2 = K$$

- P = amount of zero padding
- S = Stride
- F = Spatial extent
- K = No of Filters

$S \rightarrow$  stride movement of filters  
Step size

$S=2$        $S=1$

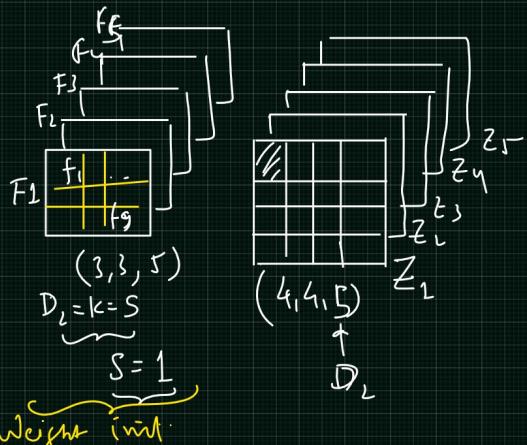
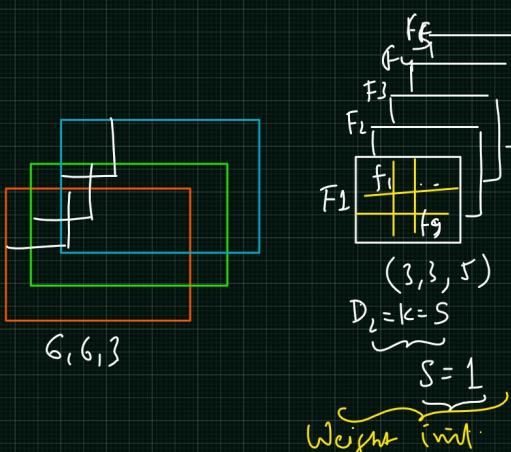
$$\text{if img} = 6, 6, 3$$

$$W_2 = \frac{6 - 3 + 0}{1} + 1 = \frac{6-3}{1} + 1 = 4$$

$$H_2 = \frac{6 - 3 + 0}{1} + 1 = \frac{6-3}{1} + 1 = 4$$

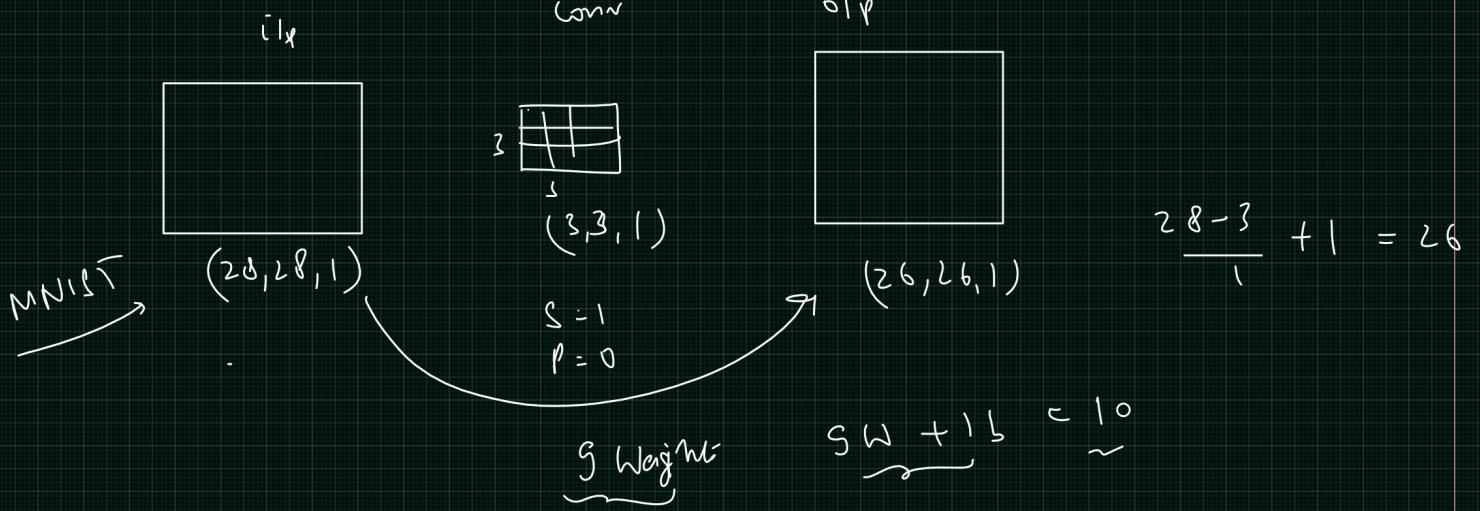
$$D_2 = K = \text{no. of filters}$$

$$(4,4,1)$$



$$RGB \otimes F1 = Z_1$$

$$\begin{aligned}
 Z_{11} = & f_1 r_1 + f_2 r_2 + \dots \\
 & + f_1 g_1 + f_2 g_2 + \dots \\
 & + f_1 b_1 + f_2 b_2 + \dots
 \end{aligned}$$



$$\text{Input: } (L \times W) \rightarrow \underbrace{784}_{\sim} \times \underbrace{3W}_{\sim} = \underbrace{W}_{\sim}$$

$$\begin{matrix}
 \circ \\
 \circ \\
 \circ \\
 \vdots \\
 784
 \end{matrix}
 \quad \underbrace{3W}_{\sim}$$

$$W = 784 \times 3W + b$$

Input:  $I$   
 $x_{11}, x_{12}, x_{13}$  highlighted in red

|          |          |          |  |          |
|----------|----------|----------|--|----------|
| $x_{11}$ | $x_{12}$ | $x_{13}$ |  |          |
| $x_{11}$ | $x_{12}$ | $x_{13}$ |  |          |
| $x_{31}$ | $x_{32}$ | $x_{33}$ |  |          |
|          |          |          |  | $x_{66}$ |

Feature map:  $F$   
 $w_{11}, w_{12}, w_{13}$  highlighted in green

|          |          |          |
|----------|----------|----------|
| $w_{11}$ | $w_{12}$ | $w_{13}$ |
| $w_{21}$ | $w_{22}$ | $w_{23}$ |
| $w_{31}$ | $w_{32}$ | $w_{33}$ |

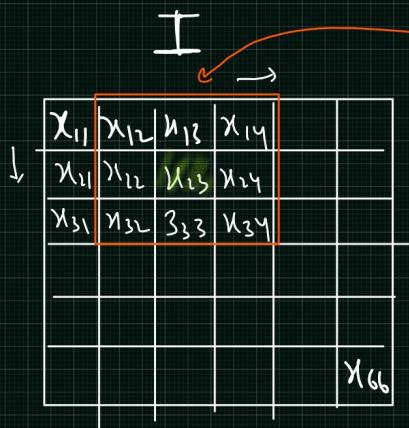
$(3, 3, 1)$

Output:  $Z_{11}$   
 $Z_{44}$

$(4, 4, 1)$

$$(6, 6, 1)$$

$$\begin{aligned}
 Z_{11} &= w_{11}x_{11} + w_{12}x_{12} + w_{13}x_{13} + w_{21}x_{21} + w_{22}x_{22} + \dots + b \quad \text{bias} \\
 &= \begin{bmatrix} w_{11} \\ w_{12} \\ \vdots \\ w_{33} \end{bmatrix}^T \begin{bmatrix} x_{11} \\ x_{12} \\ \vdots \\ x_{33} \end{bmatrix} + b = w^T x + b
 \end{aligned}$$



$F$

$$\begin{matrix} w_{11} & w_{12} & w_{13} \\ w_{21} & w_{22} & w_{23} \\ w_{31} & w_{32} & w_{33} \end{matrix}$$

$$(3, 3, 1)$$

$\omega$

$$\begin{matrix} z_{11} & z_{12} & & \\ & & & \\ & & & z_{44} \end{matrix}$$

$(4, 4, 1)$

$$S = 1$$

$P = 0$

$$(6, 6, 1)$$

$$z_{12} = (w_{11} x_{12} + w_{12} x_{13} + \dots + w_{33} x_{34}) + \text{bias}$$

$$= W^T \cdot x + \text{bias}$$

$I$

Apply padding

$F$

$$\begin{matrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & x_{11} & x_{12} & x_{13} & x_{14} & & \\ 0 & x_{21} & x_{22} & x_{23} & x_{24} & & \\ 0 & x_{31} & x_{32} & x_{33} & x_{34} & & \\ 0 & & & & & & \\ 0 & & & & & & \\ 0 & & & & & & \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{matrix}$$

$$(6, 6, 1)$$

$$(8, 8, 1)$$

$$\begin{matrix} w_{11} & w_{12} & w_{13} \\ w_{21} & w_{22} & w_{23} \\ w_{31} & w_{32} & w_{33} \end{matrix}$$

$$(3, 3, 1)$$

$\omega$

$$\begin{matrix} z_{11} & z_{12} & & \\ & & & \\ & & & z_{44} \end{matrix}$$

$$S = 1$$

$P = 1$

$$(4, 4, 1)$$

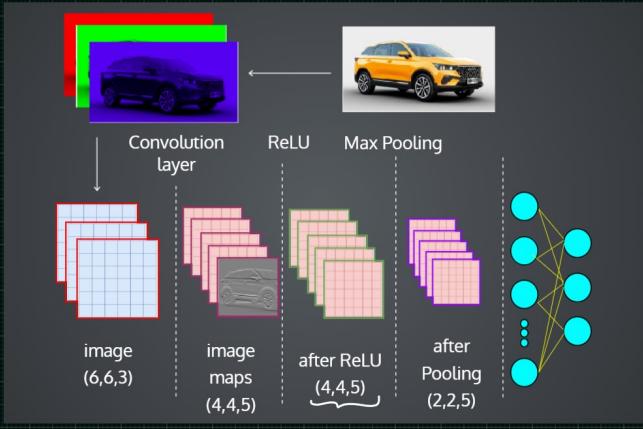
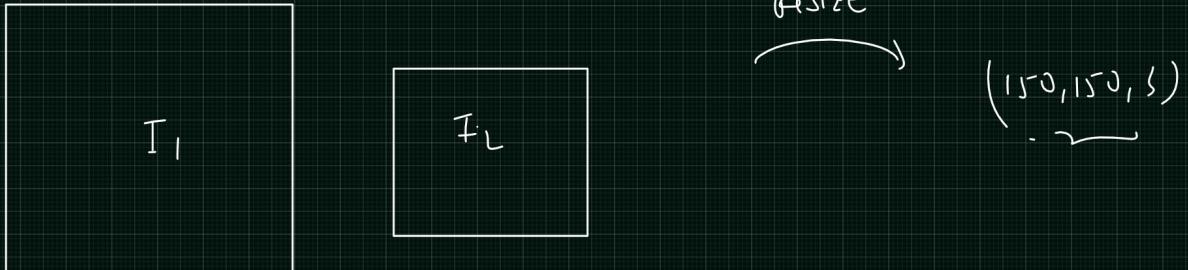
$(\underbrace{6, 6, 1})$

$$\frac{W - F + 2P + 1}{1}$$

$$\frac{6 - 3 + 2(1) + 1}{1} = 3 + 2 + 1 = 6$$

$$z_{11} = w_{11} \cdot 0 + w_{12} \cdot 0 + w_{13} \cdot 0 + w_{21} \cdot 0 + w_{22} \cdot x_{11} + w_{23} \cdot x_{12} \\ w_{31} \cdot x_0 + w_{32} \cdot x_4 + w_{33} \cdot x_{22}$$





$$\begin{matrix} L_1, L_2, L_3 \\ W_1, H_1, D_1 \end{matrix}$$

$$f = 2$$

$$S = 2$$

### Summary of Pooling Layer

$$\text{input image size} = W_1 \times H_1 \times D_1$$

$$\text{output image size} = W_2 \times H_2 \times D_2$$

$$W_2 = \frac{(W_1 - f)}{S} + 1$$

$$H_2 = \frac{(H_1 - f)}{S} + 1$$

$$D_2 = D_1$$

$$W_2 = \frac{W_1 - f}{S} + 1$$

$$= \frac{4 - 2}{2} + 1 = \frac{2}{2} + 1 = 1 + 1 = 2$$

$$H_2 = 2$$

$$D_2 = D_1 \Rightarrow (2, 2, 5)$$

$$2, 2, 5 = 2 \times 2 \times 5 = 20 \rightarrow$$

$$\begin{matrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{matrix}$$

20 units -  
AVG

filter weights  $\rightarrow$

$$\frac{\partial C}{\partial W_{11}}$$

|             |          |          |          |
|-------------|----------|----------|----------|
| $z_{11}$    | $z_{12}$ | $z_{13}$ | $z_{14}$ |
| $z_{21}$    | $z_{22}$ | $z_{23}$ | $z_{24}$ |
| $z_{44}$    |          |          |          |
| $(4, 4, 1)$ |          |          |          |

Max pooling

$(2, 2)$

$y_{11}$

$S = 2$

$$y_{11} = \max(z_{11}, z_{12}, z_{21}, z_{22})$$

|             |          |          |          |
|-------------|----------|----------|----------|
| $z_{11}$    | $z_{12}$ | $z_{13}$ | $z_{14}$ |
| $z_{21}$    | $z_{22}$ | $z_{23}$ | $z_{24}$ |
| $z_{44}$    |          |          |          |
| $(4, 4, 1)$ |          |          |          |

Max pooling

$(2, 2)$

$S = 2$

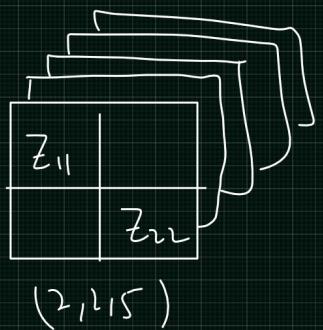
|          |          |
|----------|----------|
| $y_{11}$ | $y_{12}$ |
| $y_{21}$ | $y_{22}$ |

$(2, 2, 1)$

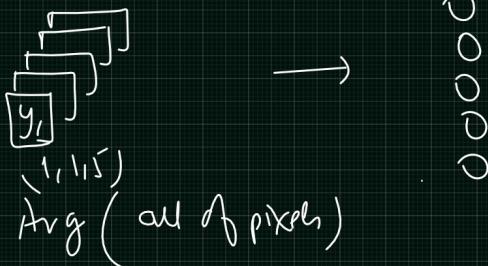
$$y_{11} = \max(z_{11}, z_{12}, z_{21}, z_{22})$$

$$y_{12} = \max(z_{13}, z_{14}, z_{23}, z_{24})$$

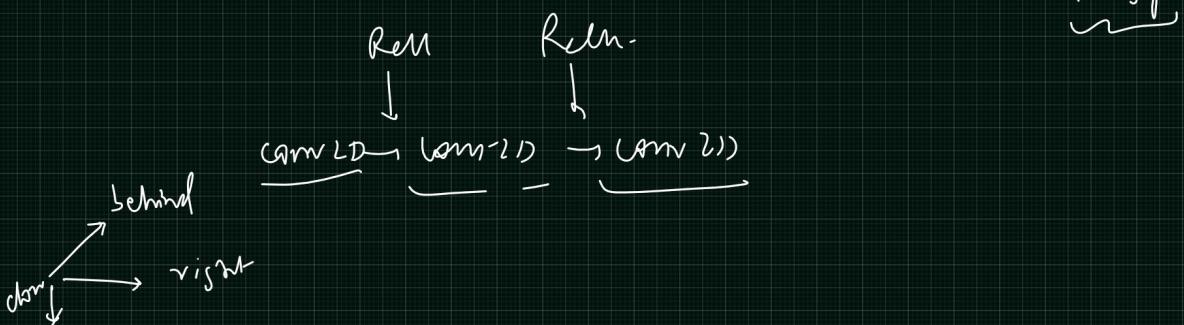
Global Avg. Pooling (Alternative to flattening)

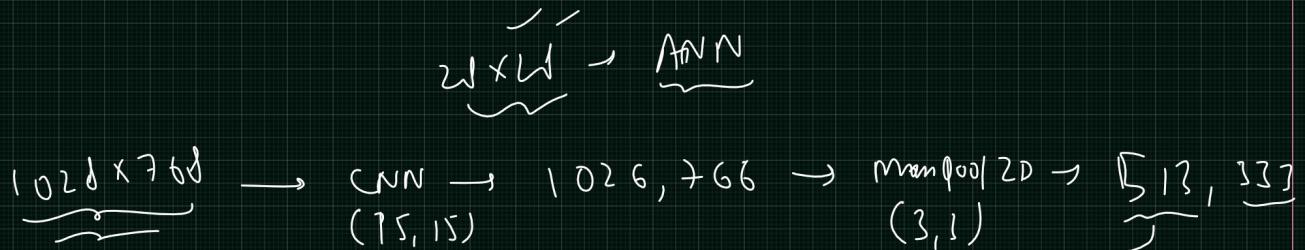
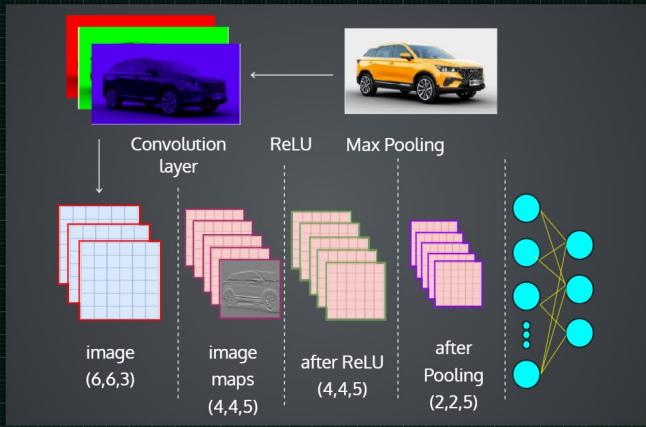
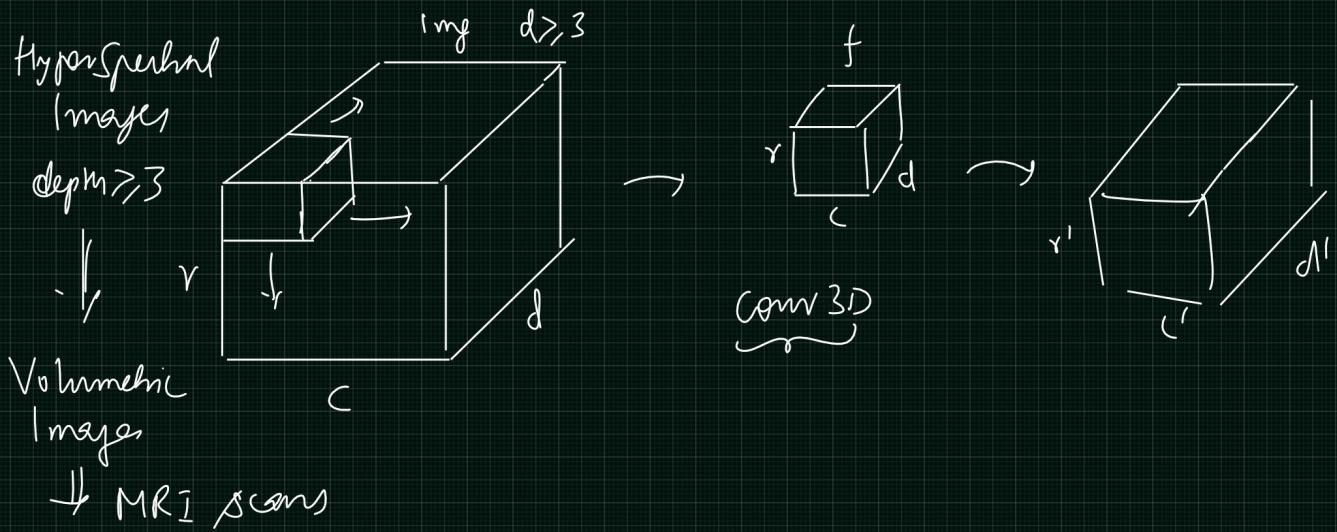


$(2, 2, 1)$



$$y_1 = \text{Avg}(\text{all of pixels})$$





$\xrightarrow{\text{CNN - } \xrightarrow{(32, 20)}}$

$\begin{matrix} C \\ \text{Conv} \\ \text{Conv} \end{matrix} \xrightarrow{\text{Conv}} \begin{matrix} P_{\text{ReLU}} \\ \text{ReLU} \end{matrix} \xrightarrow{\text{Conv}}$

CNN

