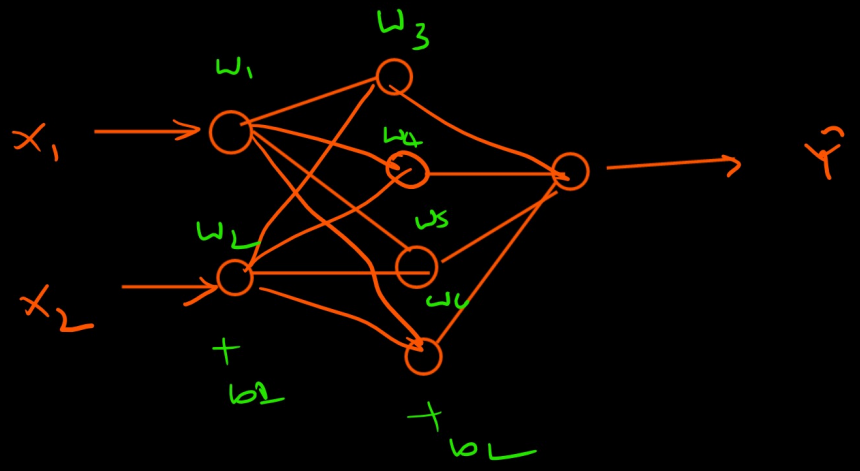
Backpropagation in CNN

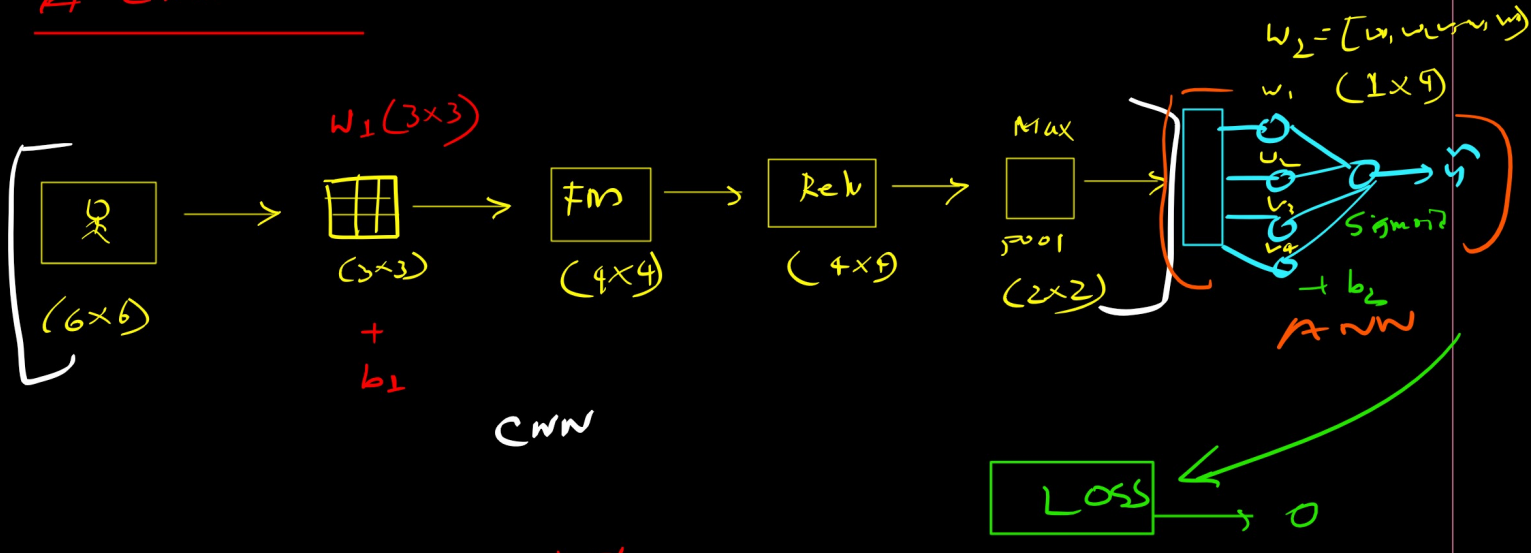
ANN →



$W = w_1, w_2, w_3, w_4, w_5, w_6$

$b = b_1, b_2$

CNN →



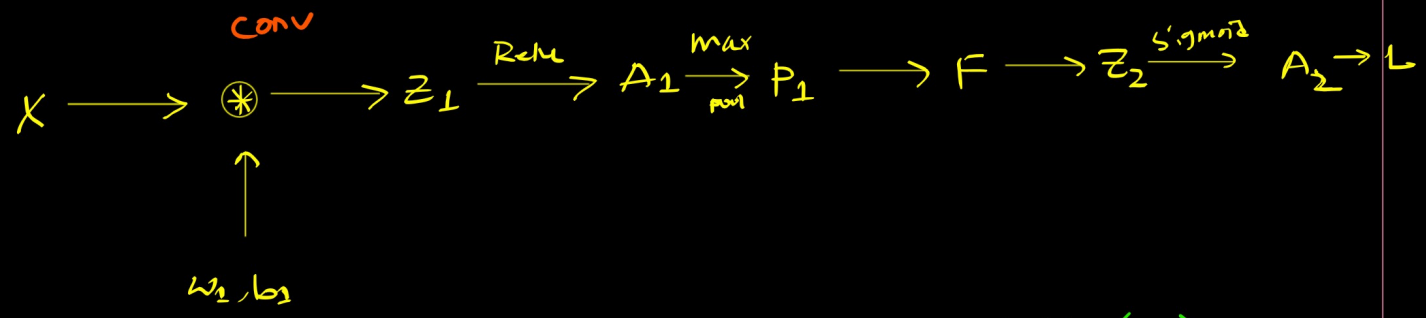
Trainable parameters

$W_1 = (3\times3)$ $\quad W_2 = (1\times4)$ $\quad \Rightarrow$ 15 Trainable parameters

$b_1 = (1\times1)$ $\quad b_2 = (1,1)$

$$L = -y_i \log(\hat{y}_i) - (1-y_i)\log(1-\hat{y}_i)$$

Logical flow:



$$Z_1 = \text{conv}((X, W_1) + b_1)$$

$$A_1 = \text{Relu}(Z_1)$$

$$P_1 = \text{maxpool}(A_1)$$

$$F = \text{Flatten}(P_1)$$

$$Z_2 = (W_2 \cdot F) + b_2$$

$$A_2 = \sigma(Z_2)$$

Forward pass → Gradient Descent

$$W_1 = W_1 - \eta \frac{\partial L}{\partial W_1}$$

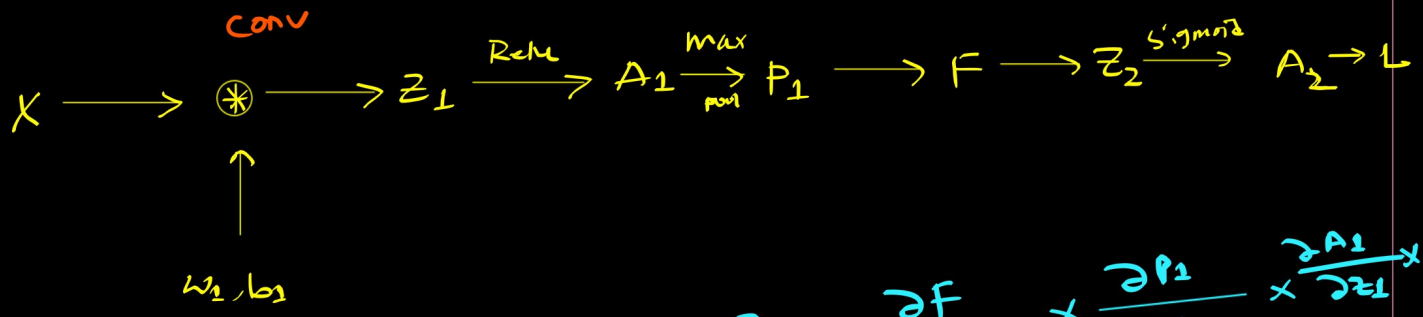
$$b_1 = b_1 - \eta \frac{\partial L}{\partial b_1}$$

$$W_2 = W_2 - \eta \frac{\partial L}{\partial W_2}$$

$$b_2 = b_2 - \eta \frac{\partial L}{\partial b_2}$$

Backward pass:

Logical flow:



$$\frac{\partial L}{\partial W_1} = \frac{\partial L}{\partial A_2} \times \frac{\partial A_2}{\partial Z_2} \times \frac{\partial Z_2}{\partial F} \times \frac{\partial F}{\partial P_1} \times \frac{\partial P_1}{\partial A_1} \times \frac{\partial A_1}{\partial Z_1} \times \frac{\partial Z_1}{\partial W_1}$$

$$\frac{\partial L}{\partial b_1} = \frac{\partial L}{\partial A_2} \times \frac{\partial A_2}{\partial Z_2} \times \frac{\partial Z_2}{\partial F} \times \frac{\partial F}{\partial P_1} \times \frac{\partial P_1}{\partial A_1} \times \frac{\partial A_1}{\partial Z_1} \times \frac{\partial Z_1}{\partial b_1}$$

$$\frac{\partial L}{\partial W_2} = \frac{\partial L}{\partial A_2} \times \frac{\partial A_2}{\partial Z_2} \times \frac{\partial Z_2}{\partial W_2}$$

$$\frac{\partial L}{\partial b_2} = \frac{\partial L}{\partial A_2} \times \frac{\partial A_2}{\partial Z_2} \times \frac{\partial Z_2}{\partial b_2}$$

Loss $\longrightarrow$ 0

Backpropagation in CNN

Transfer learning → pretrained