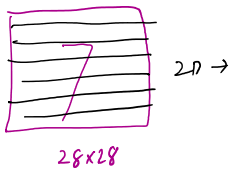
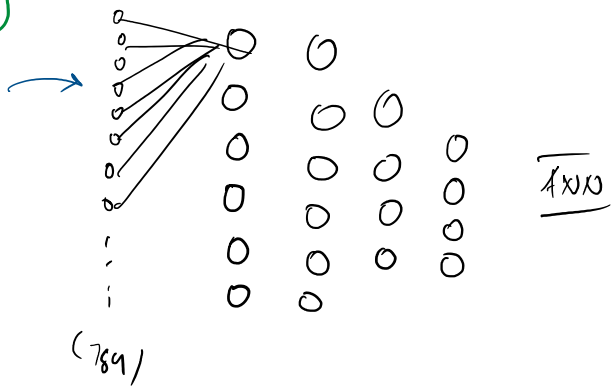
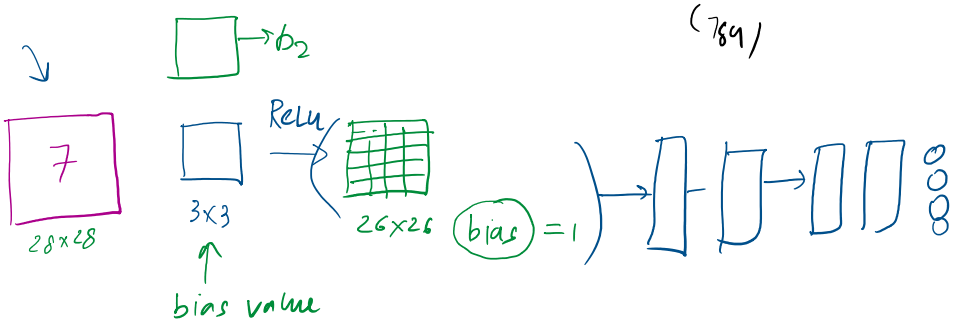
# CNN Vs ANN

06 September 2022 10:00

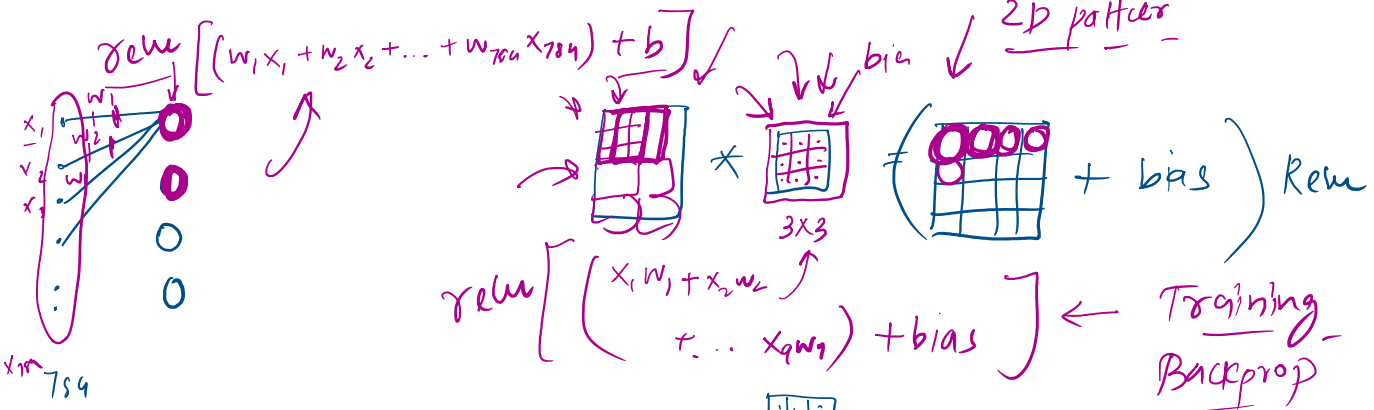
1) Computation Cost → ↓
2) Overfitting →
3) Loss of imp features like spatial arrangement of pixels

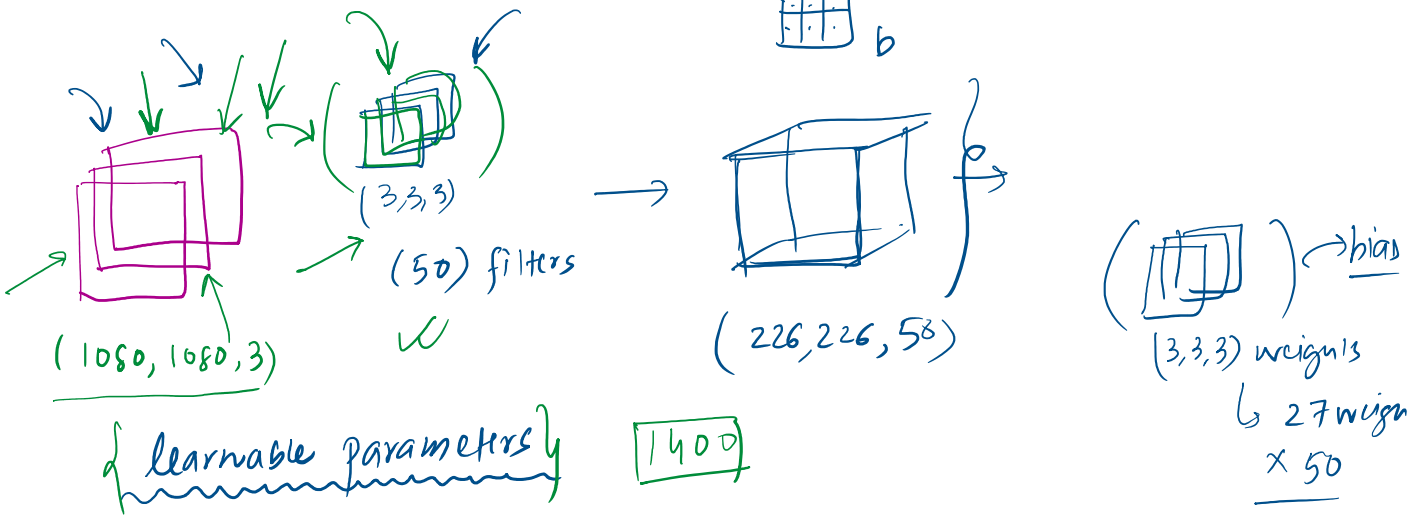


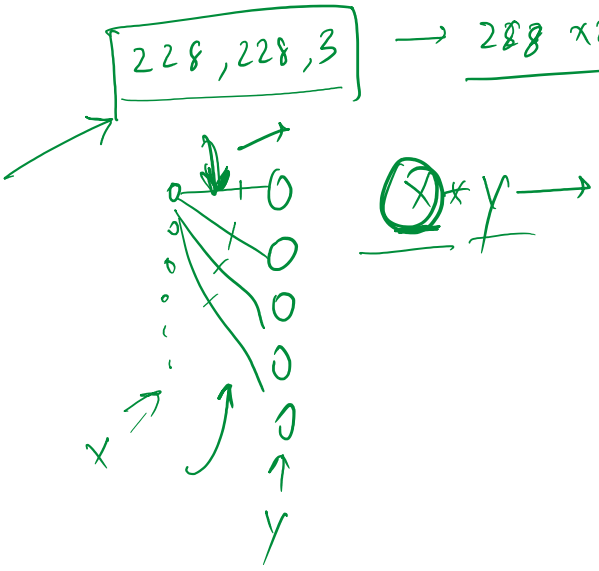




similarity ANN CNN




learnable parameters 1400




288 x228x3

1080 X 1080 X 3

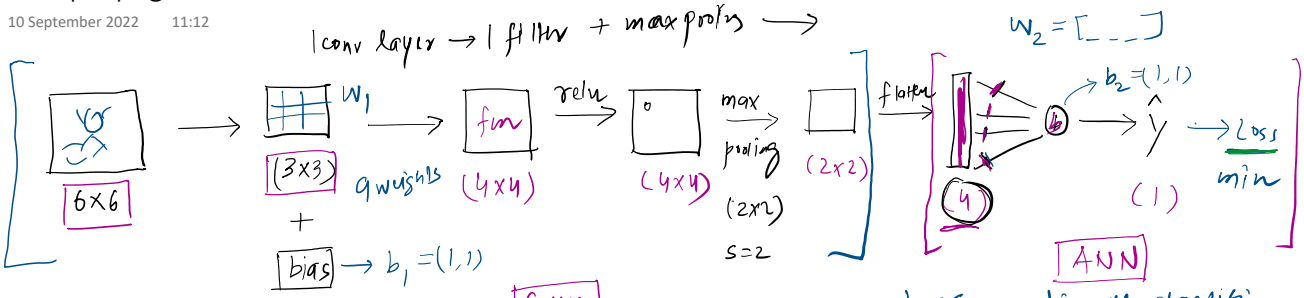
overfitting AND

Backprop In CNN

1350 wts
+50
1400

# Backpropagation in CNN

10 September 2022 11:12




Loss → binary classif

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

$A_2 = \hat{y}$

Trainable Parameters

$W_1 = (3,3)$ $W_2 = (1,4)$ $= 15$ trainable parameters

$b_1 = (1,1)$ $b_2 = (1,1)$

Logical Flow

Forward Prop

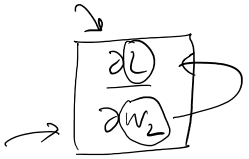
$$\begin{cases} Z_1 = \text{conv}(X, W_1) + b_1 \\ A_1 = \text{relu}(Z_1) \\ P_1 = \text{maxpool}(A_1) \end{cases} \qquad \begin{cases} F = \text{flatten}(P_1) \\ Z_2 = W_2 F + b_2 \\ A_2 = \sigma(Z_2) \end{cases}$$

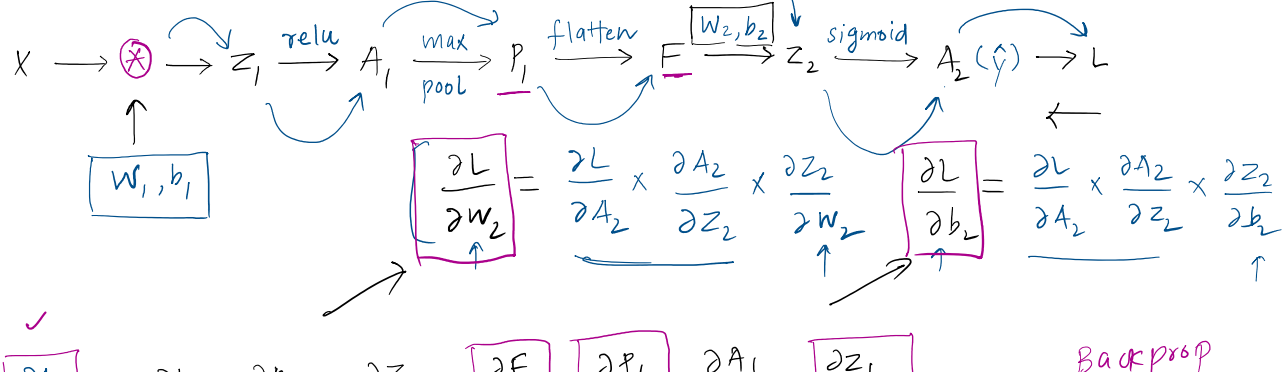
Gradient Descent

$W_1 = W_1 - \eta \frac{\partial L}{\partial W_1}$ ✓ $\quad W_2 = W_2 - \eta \frac{\partial L}{\partial W_2}$ ✓

$b_1 = b_1 - \eta \frac{\partial L}{\partial b_1}$ ✓ $\quad b_2 = b_2 - \eta \frac{\partial L}{\partial b_2}$ ✓

Loss is minimized



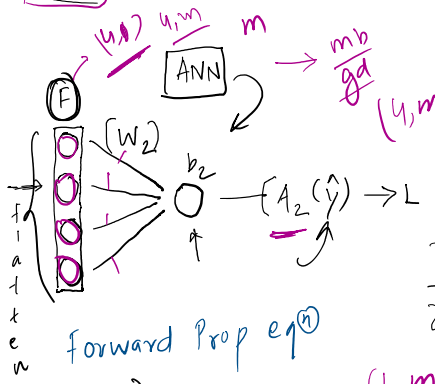




Backprop

$$\frac{\partial L}{\partial w_1} = \frac{\partial L}{\partial A_2} \times \frac{\partial A_2}{\partial Z_2} \times \frac{\partial Z_2}{\partial F} \times \frac{\partial F}{\partial P_1} \times \frac{\partial P_1}{\partial A_1} \times \frac{\partial A_1}{\partial Z_1} \times \frac{\partial Z_1}{\partial w_1}$$

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

Backprop
- → Convolution
- → Flatten
- → Max pooling




$$\frac{\partial L}{\partial w_2} = \frac{\partial L}{\partial A_2} \frac{\partial A_2}{\partial Z_2} \frac{\partial Z_2}{\partial w_2} \qquad \frac{\partial L}{\partial b_2} = \frac{\partial L}{\partial A_2} \frac{\partial A_2}{\partial Z_2} \frac{\partial Z_2}{\partial b_2}$$

Forward Prop eqⁿ

$$\begin{cases} Z_2 = W_2 F + b_2 \\ A_2 = \sigma(Z_2) \end{cases}$$

$$\frac{\partial L}{\partial a_2} = \frac{\partial}{\partial a_2}\left[-y_i \log(a_2) - (1-y_i)\log(1-a_2)\right]$$

$$(1,m) \quad Y = (1,m) \qquad = -\frac{y_i}{a_2} + \frac{(1-y_i)}{(1-a_2)} = \frac{-y_i(1-a_2) + a_2(1-y_i)}{a_2(1-a_2)}$$

$$\frac{\partial L}{\partial a_2} = \frac{-y_i + \cancel{y_i a_2} + a_2 - \cancel{a_2 y_i}}{a_2(1-a_2)} = \frac{(a_2 - y_i)}{a_2(1-a_2)}$$

$$\left[\frac{\partial A_2}{\partial Z_2} = \sigma(Z_2)[1-\sigma(Z_2)] = a_2[1-a_2]\right]$$

$W_2$ update

Shape =

$$\boxed{\frac{\partial Z_2}{\partial w_2} = F} \qquad \frac{\partial Z_2}{\partial b_2} = 1$$

$$\boxed{\frac{\partial L}{\partial w_2}} = \frac{(a_2 - y_i)}{\cancel{a_2(1-a_2)}} \times \cancel{a_2(1-a_2)} \times F = (a_2 - y_i)F = (A_2 - Y)F^T$$

(1,4) (4,1) (1,1) (1,1) (1,4) (1,4)

$$\frac{\partial L}{\partial b_2} = \frac{(a_2 - y_i)}{\cancel{a_2(1-a_2)}} \cancel{a_2(1-a_2)} \times 1 = (A_2 - Y)$$

m images

$$\boxed{\frac{\partial L}{\partial w_2} = (A_2 - Y)F^T} \qquad \boxed{\frac{\partial L}{\partial b_2} = (A_2 - Y)}$$

$\frac{\partial L}{\partial w_2} = (A_2 - Y)A_1^T$ | $\frac{\partial L}{\partial b_2} =$

$(1,m) - (1,m)$

$(1,m)\ (m,4) \rightarrow (1,4)$

1 image

baken of image

$W_2 \rightarrow (1,4)$