CNN

ECE C147/247

1. CNN forward pass — Create output matrix (What is the shape?)

Input shape : (N, C, H, W) Filter shape (F, C, H_f, W_f)

Output shape : (N, F, H', W')

$$H' = \frac{H - H_f + 2pad}{stride} + 1 \qquad W' = \frac{W - W_f + 2pad}{stride} + 1$$ Lecture note 11

Output = np.zeros (...)

```
for n in range(N):
    for i range(H'):
        for j range(W'):
```

np. sum(W × X_{seg}, axis=( )) + b

Output [n, :, i, j] = Segment of the input × filter + b

Find index for x_seg :

Find padded input : x → x_pad , np.pad

Use i and j to find x_seg ⟹ use stride

2. Backward pass for a CNN layer ⟹ just like another CNN forward

First see a 1D example :

Input $x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}$ Filter $W = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}$ Bias : b

Output $y = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$ (zero padding , stride 1)

$$\begin{cases} y_1 = w_1 x_1 + w_2 x_2 + b \\ y_2 = w_1 x_2 + w_2 x_3 + b \\ y_3 = w_1 x_3 + w_2 x_4 + b \end{cases}$$

Backprop: Input $x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}$ Filter $W = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}$ Output $y = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$

$$\frac{\partial L}{\partial y} = \begin{bmatrix} \frac{\partial L}{\partial y_1} \\ \frac{\partial L}{\partial y_2} \\ \frac{\partial L}{\partial y_3} \end{bmatrix} \qquad \begin{cases} y_1 = w_1 x_1 + w_2 x_2 + b \\ y_2 = w_1 x_2 + w_2 x_3 + b \\ y_3 = w_1 x_3 + w_2 x_4 + b \end{cases}$$

dx, dw, db ?

$$db = \frac{dy}{db} \cdot \frac{dL}{dy} = \sum_{j=1}^{3} \frac{dL}{dy_j} = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} \frac{\partial L}{\partial y_1} \\ \frac{\partial L}{\partial y_2} \\ \frac{\partial L}{\partial y_3} \end{bmatrix}$$

1

$$dw = \frac{dy}{dw} \cdot \frac{dL}{dy}$$

$$\frac{dy}{dw} = \begin{bmatrix} \frac{dy_1}{dw_1} & \frac{dy_1}{dw_2} \\ \frac{dy_2}{dw_1} & \frac{dy_2}{dw} \\ \frac{dy_3}{dw_1} & \frac{dy_3}{dw_2} \end{bmatrix} = \begin{bmatrix} x_1 & x_2 \\ x_2 & x_3 \\ x_3 & x_4 \end{bmatrix}$$

$$dw_1 = x_1 \frac{\partial L}{\partial y_1} + x_2 \frac{\partial L}{\partial y_2} + x_3 \frac{\partial L}{\partial y_3} \qquad dw_2 = x_2 \frac{\partial L}{\partial y_1} + x_3 \frac{\partial L}{\partial y_2} + x_4 \frac{\partial L}{\partial y_3}$$

$$dw = \begin{bmatrix} x_1 & x_2 \\ x_2 & x_3 \\ x_3 & x_4 \end{bmatrix} \begin{bmatrix} \frac{\partial L}{\partial y_1} \\ \frac{\partial L}{\partial y_2} \\ \frac{\partial L}{\partial y_3} \end{bmatrix}$$

$$dx = \frac{\partial y}{\partial x} \cdot \frac{\partial L}{\partial y}$$

$$\frac{\partial y}{\partial x} = \begin{bmatrix} \frac{\partial y_1}{\partial x_1} & \frac{\partial y_1}{\partial x_2} & \frac{\partial y_1}{\partial x_3} & \frac{\partial y_1}{\partial x_4} \\ \frac{\partial y_2}{\partial x_1} & \frac{\partial y_2}{\partial x_2} & \frac{\partial y_2}{\partial x_3} & \frac{\partial y_2}{\partial x_4} \\ \frac{\partial y_3}{\partial x_1} & \frac{\partial y_3}{\partial x_2} & \frac{\partial y_3}{\partial x_3} & \frac{\partial y_3}{\partial x_4} \end{bmatrix} = \begin{bmatrix} w_1 & w_2 & 0 & 0 \\ 0 & w_1 & w_2 & 0 \\ 0 & 0 & w_1 & w_2 \end{bmatrix}$$

$$= \begin{bmatrix} w_1 & w_2 & 0 & 0 \\ 0 & w_1 & w_2 & 0 \\ 0 & 0 & w_1 & w_2 \end{bmatrix} \begin{bmatrix} \frac{\partial L}{\partial y_1} \\ \frac{\partial L}{\partial y_2} \\ \frac{\partial L}{\partial y_3} \end{bmatrix}$$

$$= \begin{bmatrix} 0 \\ dy_1 \\ dy_2 \\ dy_3 \\ 0 \end{bmatrix} * \begin{bmatrix} w_2 \\ w_1 \end{bmatrix}$$

$$dx_1 = \frac{\partial y_1}{\partial x_1} \cdot \frac{\partial L}{\partial y_1} = w_1 \frac{\partial L}{\partial y_1}$$

$$dx_2 = \frac{\partial y_1}{\partial x_2} \cdot \frac{\partial L}{\partial y_1} + \frac{\partial y_2}{\partial x_1} \frac{\partial L}{\partial y_2} = w_2 \frac{\partial L}{\partial y_1} + w_1 \frac{\partial L}{\partial y_2}$$

$$dx_3 = \frac{\partial y_2}{\partial x_3} \frac{\partial L}{\partial y_2} + \frac{\partial y_3}{\partial x_3} \frac{\partial L}{\partial y_3} = w_2 \frac{\partial L}{\partial y_2} + w_3 \frac{\partial L}{\partial y_3}$$

$$dx_4 = \frac{\partial y_3}{\partial x_4} \frac{\partial L}{\partial y_3} = w_2 \frac{\partial L}{\partial y_3}$$

2D example

Input $x = \begin{bmatrix} x_{11} & x_{12} & x_{13} & x_{14} \\ x_{21} & x_{22} & x_{23} & x_{24} \\ x_{31} & x_{32} & x_{33} & x_{34} \\ x_{41} & x_{42} & x_{43} & x_{44} \end{bmatrix}$ Filter $W = \begin{bmatrix} w_{11} & w_{12} \\ w_{21} & w_{22} \end{bmatrix}$

No pading, stride 1

Output $y = \begin{bmatrix} y_{11} & y_{12} & y_{13} \\ y_{21} & y_{22} & y_{23} \\ y_{31} & y_{32} & y_{33} \end{bmatrix}$

$$\begin{cases} y_{11} = W_{11} \cdot x_{11} + W_{12} x_{12} + W_{21} x_{21} + W_{22} x_{22} + b \\ \vdots \\ y_{12} = W_{11} x_{12} + W_{12} x_{13} + W_{21} x_{22} + W_{22} x_{23} + b \\ \vdots \\ y_{21} = W_{11} x_{21} + W_{12} x_{22} + W_{21} x_{31} + W_{22} x_{32} + b \\ \vdots \end{cases}$$

Forward pass:

$$y_{ij} = \left[\sum_{k=1}^{2} \sum_{l=1}^{2} W_{kl} \cdot (x_{i+k-1,\, j+l-1})\right] + b$$

$$db = \overset{1}{\frac{\partial y}{\partial b}} \cdot \frac{\partial L}{\partial y} = \sum_{i=1}^{3} \sum_{j=1}^{3} \frac{\partial L}{\partial y_{ij}}$$

$$\frac{\partial L}{\partial y} = \begin{bmatrix} \overset{\frac{\partial L}{\partial y_{11}}}{dy_{11}} & dy_{12} & dy_{13} \\ dy_{21} & dy_{22} & dy_{23} \\ dy_{31} & dy_{32} & dy_{33} \end{bmatrix}$$

$$dW_{mn} = \sum_{i=1}^{3} \sum_{j=1}^{3} \frac{\partial y_{ij}}{\partial W_{mn}} \cdot \frac{\partial L}{\partial y_{ij}}$$

$$= \sum_{i=1}^{3} \sum_{j=1}^{3} dy_{ij} \cdot (x_{i+m-1,\, j+n-1})$$

$$\frac{\partial y_{ij}}{\partial W_{mn}} = \sum_{k=1}^{2} \sum_{l=1}^{2} \frac{\partial W_{kl}}{\partial W_{mn}} x_{i+k-1,\, j+l-1} = x_{i+m-1,\, j+n-1}$$

$$\frac{\partial W_{kl}}{\partial W_{mn}} = \begin{cases} 1 & \text{only when } k=m,\ l=n \\ 0 & \text{otherwise} \end{cases}$$

$$dW = \underbrace{\begin{bmatrix} x_{11} & x_{12} & x_{13} & x_{14} \\ x_{21} & x_{22} & x_{23} & x_{24} \\ x_{31} & x_{32} & x_{33} & x_{34} \\ x_{41} & x_{42} & x_{43} & x_{44} \end{bmatrix}}_{\text{Input}} \underbrace{\begin{bmatrix} dy_{11} & dy_{12} & dy_{13} \\ dy_{21} & dy_{22} & dy_{23} \\ dy_{31} & dy_{32} & dy_{33} \end{bmatrix}}_{\text{filter}} = x * dy$$

$$dW_{12} = x_{12} dy_{11} + x_{13} dy_{12} + x_{14} dy_{13} + x_{22} dy_{21} + x_{23} dy_{22} + x_{24} dy_{23} + x_{32} dy_{31} + x_{33} dy_{32} + x_{34} dy_{33}$$

$$dx_{mn} = \sum_{i=1}^{3}\sum_{j=1}^{3} \frac{\partial y_{ij}}{\partial x_{mn}} \frac{\partial L}{\partial y_{ij}}$$

$$\frac{\partial y_{ij}}{\partial x_{mn}} = \sum_{k=1}^{2}\sum_{l=1}^{2} w_{kl} \cdot \boxed{\frac{\partial x_{i+k-1,\, j+l-1}}{\partial x_{mn}}} = \begin{cases} 1 & \text{if } m = i+k-1 \\ & \quad n = j+l-1 \\ 0 & \text{otherwise} \end{cases}$$

$$\frac{\partial y_{ij}}{\partial x_{mn}} = W_{m-i+1,\, n-j+1}$$

$$\Rightarrow k = m-i+1 \ , \ l = n-j+1$$

$$\frac{\partial L}{\partial x_{mn}} = \sum_{i=1}^{3}\sum_{j=1}^{3} (W_{m-i+1,\, n-j+1}) \cdot dy_{ij}$$

$$dx = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & dy_{11} & dy_{12} & dy_{13} & 0 \\ 0 & dy_{21} & dy_{22} & dy_{23} & 0 \\ 0 & dy_{31} & dy_{32} & dy_{33} & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} * \begin{bmatrix} W_{22} & W_{21} \\ W_{12} & W_{11} \end{bmatrix}$$

Input filter

$$\text{Filter } W = \begin{bmatrix} W_{11} & W_{12} \\ W_{21} & W_{22} \end{bmatrix}$$