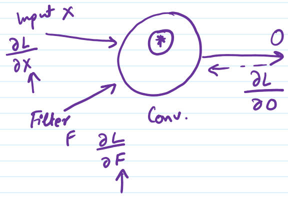
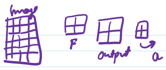


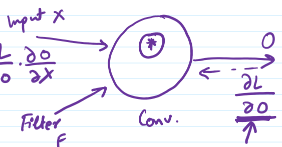


$$\left(\begin{bmatrix} X_{11} & X_{12} & X_{13} \\ X_{21} & X_{22} & X_{23} \\ X_{31} & X_{32} & X_{33} \end{bmatrix} \circledast \begin{bmatrix} F_{11} & F_{12} \\ F_{21} & F_{22} \end{bmatrix} \right) = \begin{bmatrix} O_{11} & O_{12} \\ O_{21} & O_{22} \end{bmatrix}$$

Input X — Filter F — Output

$$O_{11} = X_{11}F_{11} + X_{12}F_{22} + X_{21}F_{21} + X_{22}F_{22}$$
$$O_{12} = X_{12}F_{11} + X_{13}F_{12} + X_{22}F_{21} + X_{23}F_{22}$$
$$O_{21} = X_{21}F_{11} + X_{22}F_{12} + X_{31}F_{21} + X_{32}F_{22}$$
$$O_{22} = X_{22}F_{11} + X_{23}F_{12} + X_{32}F_{21} + X_{33}F_{22}$$

$X_{new} = X_{old} - \alpha \frac{\partial L}{\partial X}$ $\qquad \frac{\partial L}{\partial X} = \frac{\partial L}{\partial O} \cdot \frac{\partial O}{\partial X}$



$F_{new} = F_{old} - \alpha \frac{\partial L}{\partial F}$ $\qquad \frac{\partial L}{\partial F} = \frac{\partial L}{\partial O} \cdot \frac{\partial O}{\partial L}$

(1) $\frac{\partial L}{\partial F} = \frac{\partial L}{\partial O} \cdot \frac{\partial O}{\partial F}$

already known
(Loss Gradient from next layer)

$\frac{\partial O}{\partial F}$ $\qquad O = \begin{bmatrix} O_{11} & O_{12} \\ O_{21} & O_{22} \end{bmatrix}$ $\qquad F = \begin{bmatrix} F_{11} & F_{12} \\ F_{21} & F_{22} \end{bmatrix}$

$$O_{11} = X_{11}F_{11} + X_{12}F_{22} + X_{21}F_{21} + X_{22}F_{22}$$

$\frac{\partial O_{11}}{\partial F_{11}} = X_{11}$, $\frac{\partial O_{11}}{\partial F_{12}} = X_{12}$, $\frac{\partial O_{11}}{\partial F_{21}} = X_{21}$, $\frac{\partial O_{11}}{\partial F_{22}} = X_{22}$ $\rightarrow$ Equation (1)

4x4

$$O_{12} = X_{12}F_{11} + X_{13}F_{12} + X_{22}F_{21} + X_{23}F_{22}$$
$$O_{21} = X_{21}F_{11} + X_{22}F_{12} + X_{31}F_{21} + X_{32}F_{22}$$
$$O_{22} = X_{22}F_{11} + X_{23}F_{12} + X_{32}F_{21} + X_{33}F_{22}$$

IIly for O_{12}, O_{21}, O_{22}

$$\begin{bmatrix} F_{11} & F_{12} \\ F_{21} & F_{22} \end{bmatrix}$$

$\frac{\partial L}{\partial F}$

$$\frac{\partial L}{\partial F_{11}} = \frac{\partial L}{\partial O_{11}} \times \frac{\partial O_{11}}{\partial F_{11}} + \frac{\partial L}{\partial O_{12}} \times \frac{\partial O_{12}}{\partial F_{11}} + \frac{\partial L}{\partial O_{21}} \frac{\partial O_{21}}{\partial F_{11}} + \frac{\partial L}{\partial O_{22}} \frac{\partial O_{22}}{\partial F_{11}}$$

$$\frac{\partial L}{\partial F_{12}} = \frac{\partial L}{\partial O_{11}} \times \frac{\partial O_{11}}{\partial F_{12}} + \frac{\partial L}{\partial O_{12}} \times \frac{\partial O_{12}}{\partial F_{12}} + \frac{\partial L}{\partial O_{21}} \frac{\partial O_{21}}{\partial F_{12}} + \frac{\partial L}{\partial O_{22}} \frac{\partial O_{22}}{\partial F_{12}}$$

$$\frac{\partial L}{\partial F_{21}} = \frac{\partial L}{\partial O_{11}} \times \frac{\partial O_{11}}{\partial F_{21}} + \frac{\partial L}{\partial O_{12}} \times \frac{\partial O_{12}}{\partial F_{21}} + \frac{\partial L}{\partial O_{21}} \frac{\partial O_{21}}{\partial F_{21}} + \frac{\partial L}{\partial O_{22}} \frac{\partial O_{22}}{\partial F_{21}}$$

$$\frac{\partial L}{\partial F_{22}} = \frac{\partial L}{\partial O_{11}} \times \frac{\partial O_{11}}{\partial F_{22}} + \frac{\partial L}{\partial O_{12}} \times \frac{\partial O_{12}}{\partial F_{22}} + \frac{\partial L}{\partial O_{21}} \frac{\partial O_{21}}{\partial F_{22}} + \frac{\partial L}{\partial O_{22}} \frac{\partial O_{22}}{\partial F_{22}}$$

Using (1)

$$\frac{\partial L}{\partial F_{11}} = \frac{\partial L}{\partial O_{11}} \times X_{11} + \frac{\partial L}{\partial O_{12}} \times X_{12} + \frac{\partial L}{\partial O_{21}} X_{21} + \frac{\partial L}{\partial O_{22}} X_{2}$$

$$\frac{\partial L}{\partial F_{12}} = \frac{\partial L}{\partial O_{11}} \times X_{11} + \frac{\partial L}{\partial O_{12}} \times X_{12} + \frac{\partial L}{\partial O_{21}} X_{21} + \frac{\partial L}{\partial O_{22}} X$$

$$\frac{\partial L}{\partial F_{21}} = \frac{\partial L}{\partial O_{11}} \times X_{11} + \frac{\partial L}{\partial O_{12}} X_{12} + \frac{\partial L}{\partial O_{21}} X_{21} + \frac{\partial L}{\partial O_{22}} X_{2}$$

$$\frac{\partial L}{\partial F_{22}} = \frac{\partial L}{\partial O_{11}} X_{11} + \frac{\partial L}{\partial O_{12}} X_{12} + \frac{\partial L}{\partial O_{21}} X_{21} + \frac{\partial L}{\partial O_{22}} X_{2}$$

$$\begin{bmatrix} \frac{\partial L}{\partial F_{11}} & \frac{\partial L}{\partial F_{12}} \\ \frac{\partial L}{\partial F_{21}} & \frac{\partial L}{\partial F_{22}} \end{bmatrix} = \text{Convolution}\left(\underbrace{\begin{bmatrix} X_{11} & X_{12} & X_{23} \\ X_{21} & X_{22} & X_{23} \\ X_{31} & X_{32} & X_{33} \end{bmatrix}}_{\text{Input } X} , \right)$$

(2) $$\frac{\partial L}{\partial X} = \underbrace{\frac{\partial L}{\partial O}} \cdot \underbrace{\frac{\partial O}{\partial X}}$$

$$O = \begin{bmatrix} O_{11} & O_{12} \\ O_{21} & O_{22} \end{bmatrix} \qquad X = \begin{bmatrix} \; & \; & \; \\ \; & \; & \; \\ \; & \; & \; \end{bmatrix} \; 3 \times 3.$$

$$O_{11} = X_{11} F_{11} + X_{12} F_{22} + X_{21} F_{21} + X_{22} F_{22}$$

$$O_{12} = X_{12} F_{11} + X_{13} F_{12} + X_{22} F_{21} + X_{23} F_{22}$$

$$O_{21} = X_{21} F_{11} + X_{22} F_{12} + X_{31} F_{21} + X_{32} F_{22}$$

$$O_{22} = X_{22} F_{11} + X_{23} F_{12} + X_{32} F_{21} + X_{33} F_{22}$$

$$\frac{\partial O_{11}}{\;} = F_{11} , \quad \frac{\partial O_{11}}{\;} = F_{12} , \quad \frac{\partial O_{11}}{\partial X} , \quad \frac{\partial O_{11}}{\partial X_{22}} = F_{22}.$$

12

21

22

2

22

2

2

$$\begin{bmatrix} \cdots & \frac{\partial L}{\partial O_{12}} \\ \cdots & \frac{\partial L}{\partial O_{22}} \end{bmatrix}$$

ns
adient
from next
layer

$$\frac{\partial O_{11}}{\partial X_{11}} = F_{11}, \quad \frac{\partial O_{11}}{\partial X_{12}} = F_{12}, \quad \frac{\partial O_{11}}{\partial X_{21}}, \quad \frac{\partial O_{11}}{\partial X_{22}} = F_{22}.$$

Illy O_{12}, O_{21}, O_{22}

$$\frac{\partial L}{\partial X} = \begin{bmatrix} \frac{\partial L}{\partial X_{11}} & \frac{\partial L}{\partial X_{12}} & \frac{\partial L}{\partial X_{13}} \\ \frac{\partial L}{\partial X_{21}} & \frac{\partial L}{\partial X_{22}} & \frac{\partial L}{\partial X_{23}} \\ \frac{\partial L}{\partial X_{31}} & \frac{\partial L}{\partial X_{32}} & \frac{\partial L}{\partial X_{33}} \end{bmatrix}$$

$$\frac{\partial L}{\partial X_{11}} = \frac{\partial L}{\partial O_{11}} \times F_{11}$$

$$\frac{\partial L}{\partial X_{12}} = \frac{\partial L}{\partial O_{11}} \times F_{12} + \frac{\partial L}{\partial O_{11}} \times F_{11}$$

$$\begin{bmatrix} \frac{\partial L}{\partial X_{11}} & & \\ & & \\ & & \frac{\partial L}{\partial X_{33}} \end{bmatrix} = \text{Full Conv.} \; ($$

$$\frac{\partial L}{\partial X_{13}} = \frac{\partial L}{\partial O_{12}} \times F_{12}$$

$$\frac{\partial L}{\partial X_{21}} = \frac{\partial L}{\partial O_{11}} F_{21} + \frac{\partial L}{\partial O_{21}} \times F_{11}$$

$$\frac{\partial L}{\partial X_{22}} = \frac{\partial L}{\partial O_{11}} \times F_{22} + \frac{\partial L}{\partial O_{12}} F_{21} + \frac{\partial L}{\partial O_{21}} \times F_{12} + \frac{\partial L}{\partial O_{22}} \times F_{11}$$

$$\frac{\partial L}{\partial X_{23}} = \frac{\partial L}{\partial O_{12}} \times F_{22} + \frac{\partial L}{\partial O_{22}} \times F_{12}$$

$$\frac{\partial L}{\partial X_{31}} = \frac{\partial L}{\partial O_{21}} \times F_{21}$$

$$\frac{\partial L}{\partial X_{32}} = \frac{\partial L}{\partial O_{21}} \times F_{22} + \frac{\partial L}{\partial O_{22}} \times F_{21}$$

$$\frac{\partial L}{\partial X_{33}} = \frac{\partial L}{\partial O_{22}} \times F_{22}$$

Both f/w & b/w pass are $\underbrace{\text{convolutions}}_{\text{of a lay}}$

$$\frac{\partial L}{\partial F} = \text{Conv.}\left(X, \frac{\partial L}{\partial O}\right)$$

$$\frac{\partial L}{} = \text{Full}\left(180^\circ, \frac{\partial L}{}\right)$$

$$\left(\underbrace{\begin{bmatrix} F_{22} & F_{21} \\ F_{12} & F_{11} \end{bmatrix}}_{180^{\circ}\ \text{Inverted}} \circledast \begin{bmatrix} \frac{\partial L}{\partial O_{11}} & \frac{\partial L}{\partial O_{12}} \\ \frac{\partial L}{\partial O_{21}} & \frac{\partial L}{\partial O_{22}} \end{bmatrix} \right)$$

s in CNN.

er

$$\frac{\partial L}{\partial X} = \underset{\text{conv.}}{\text{Full}}\left(\underset{F}{180^\circ}, \frac{\partial L}{\partial O}\right)$$