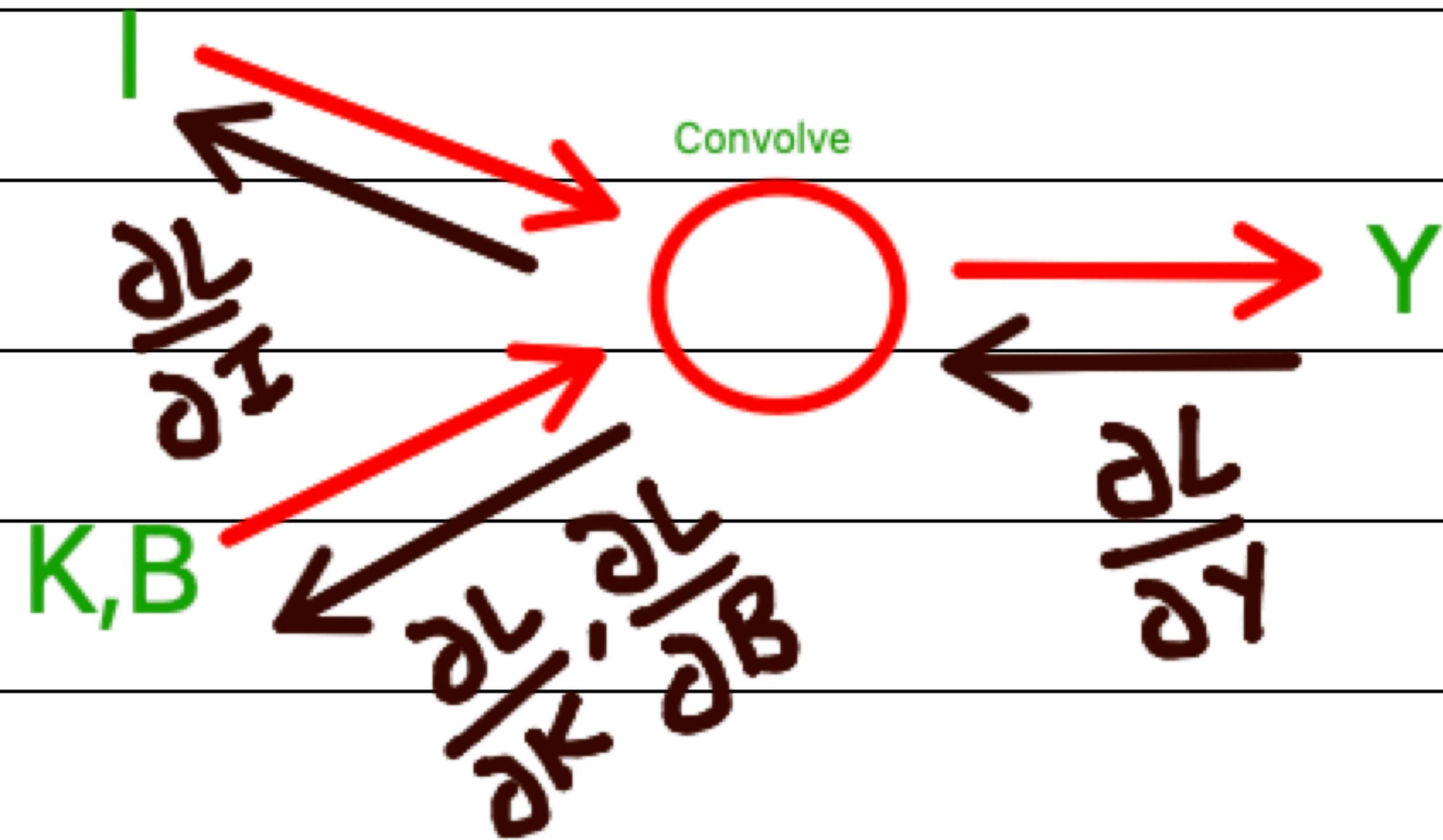


Backpropagation in CNN

SCALER



$$I = \begin{bmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \end{bmatrix}_{3 \times 3} \odot K_1 \begin{bmatrix} K_{11} & K_{12} \\ K_{21} & K_{22} \end{bmatrix}_{2 \times 2} + B_1 = Y_{2 \times 2}$$

$$Y_1 \begin{bmatrix} Y_{11} & Y_{12} \\ Y_{21} & Y_{22} \end{bmatrix}$$

$$K'_2 \begin{bmatrix} K'_{11} & K'_{12} \\ K'_{21} & K'_{22} \end{bmatrix}_{2 \times 2} + B'_2 = Y'_2 \begin{bmatrix} Y'_11 & Y'_12 \\ Y'_21 & Y'_22 \end{bmatrix}$$

$$\begin{bmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \end{bmatrix} \odot K_1 \begin{bmatrix} k_{11} & k_{12} \\ k_{21} & k_{22} \end{bmatrix} + B_1$$

SCALER 

$$y_{11} = x_{11}k_{11} + x_{12}k_{12} + x_{21}k_{21} + x_{22}k_{22} + B_1$$

$$y_{12} = x_{12}k_{11} + x_{13}k_{12} + x_{22}k_{21} + x_{23}k_{22} + B_1$$

$$y_{21} = x_{21}k_{11} + x_{22}k_{12} + x_{31}k_{21} + x_{32}k_{22} + B_1$$

$$y_{22} = x_{22}k_{11} + x_{23}k_{12} + x_{32}k_{21} + x_{33}k_{22} + B_1$$

$$\frac{\partial L}{\partial K_1} = \begin{bmatrix} \frac{\partial L}{\partial k_{11}} & \frac{\partial L}{\partial k_{12}} \\ \frac{\partial L}{\partial k_{21}} & \frac{\partial L}{\partial k_{22}} \end{bmatrix}$$

Using Chain Rule: \Rightarrow

$$\frac{\partial L}{\partial K_1} = \frac{\partial L}{\partial Y} \cdot \frac{\partial Y}{\partial K_1}$$

$$\frac{\partial L}{\partial K_{ij}} = \sum \frac{\partial L}{\partial Y_{mn}} \cdot \frac{\partial Y_{mn}}{\partial K_{ij}}$$

$$\frac{\partial L}{\partial k_{11}} = \frac{\partial L}{\partial Y_{11}} \frac{\partial Y_{11}}{\partial k_{11}} + \frac{\partial L}{\partial Y_{12}} \frac{\partial Y_{12}}{\partial k_{11}} + \frac{\partial L}{\partial Y_{21}} \frac{\partial Y_{21}}{\partial k_{11}} + \frac{\partial L}{\partial Y_{22}} \frac{\partial Y_{22}}{\partial k_{11}}$$

$$\frac{\partial L}{\partial k_{12}} = \frac{\partial L}{\partial Y_{11}} \frac{\partial Y_{11}}{\partial k_{12}} + \frac{\partial L}{\partial Y_{12}} \frac{\partial Y_{12}}{\partial k_{12}} + \frac{\partial L}{\partial Y_{21}} \frac{\partial Y_{21}}{\partial k_{12}} + \frac{\partial L}{\partial Y_{22}} \frac{\partial Y_{22}}{\partial k_{12}}$$

$$\frac{\partial L}{\partial k_{21}} = \frac{\partial L}{\partial Y_{11}} \frac{\partial Y_{11}}{\partial k_{21}} + \frac{\partial L}{\partial Y_{12}} \frac{\partial Y_{12}}{\partial k_{21}} + \frac{\partial L}{\partial Y_{21}} \frac{\partial Y_{21}}{\partial k_{21}} + \frac{\partial L}{\partial Y_{22}} \frac{\partial Y_{22}}{\partial k_{21}}$$

$$\frac{\partial L}{\partial k_{22}} = \frac{\partial L}{\partial Y_{11}} \frac{\partial Y_{11}}{\partial k_{22}} + \frac{\partial L}{\partial Y_{12}} \frac{\partial Y_{12}}{\partial k_{22}} + \frac{\partial L}{\partial Y_{21}} \frac{\partial Y_{21}}{\partial k_{22}} + \frac{\partial L}{\partial Y_{22}} \frac{\partial Y_{22}}{\partial k_{22}}$$

$$\frac{\partial L}{\partial K_i} = \begin{bmatrix} \frac{\partial L}{\partial K_{11}} & \frac{\partial L}{\partial K_{12}} \\ \frac{\partial L}{\partial K_{21}} & \frac{\partial L}{\partial K_{22}} \end{bmatrix}$$

Similarly

$$\frac{\partial L}{\partial K_{11}} = \frac{\partial L}{\partial Y_{11}} \cdot X_{11} + \frac{\partial L}{\partial Y_{12}} \cdot X_{12} + \frac{\partial L}{\partial Y_{21}} \cdot X_{21} + \frac{\partial L}{\partial Y_{22}} \cdot X_{22}$$

$$\frac{\partial L}{\partial K_{12}} = \frac{\partial L}{\partial Y_{11}} \cdot X_{12} + \frac{\partial L}{\partial Y_{12}} \cdot X_{13} + \frac{\partial L}{\partial Y_{21}} \cdot X_{22} + \frac{\partial L}{\partial Y_{22}} \cdot X_{23}$$

$$\frac{\partial L}{\partial K_{21}} = \frac{\partial L}{\partial Y_{11}} \cdot X_{21} + \frac{\partial L}{\partial Y_{12}} \cdot X_{22} + \frac{\partial L}{\partial Y_{21}} \cdot X_{31} + \frac{\partial L}{\partial Y_{22}} \cdot X_{32}$$

$$\frac{\partial L}{\partial K_{22}} = \frac{\partial L}{\partial Y_{11}} \cdot X_{22} + \frac{\partial L}{\partial Y_{12}} \cdot X_{23} + \frac{\partial L}{\partial Y_{21}} \cdot X_{32} + \frac{\partial L}{\partial Y_{22}} \cdot X_{33}$$

$$\Rightarrow \begin{bmatrix} X_{11} & X_{12} & X_{13} \\ X_{21} & X_{22} & X_{23} \\ X_{31} & X_{32} & X_{33} \end{bmatrix} \text{Conv}$$

$$\begin{bmatrix} \frac{\partial L}{\partial Y_{11}} & \frac{\partial L}{\partial Y_{12}} \\ \frac{\partial L}{\partial Y_{21}} & \frac{\partial L}{\partial Y_{22}} \end{bmatrix} =$$

$$\begin{bmatrix} \frac{\partial L}{\partial K_{11}} & \frac{\partial L}{\partial K_{12}} \\ \frac{\partial L}{\partial K_{21}} & \frac{\partial L}{\partial K_{22}} \end{bmatrix} = \frac{\partial L}{\partial K}$$

$$\frac{\partial L}{\partial K} = \text{Conv} \left(X, \frac{\partial L}{\partial Y} \right)$$

$$\frac{\partial L}{\partial \beta_1} = \frac{\partial L}{\partial Y} \cdot \frac{\partial Y}{\partial \beta_1} = \sum \frac{\partial L}{\partial Y_{mn}} \cdot \frac{\partial Y_{mn}}{\partial \beta_1}$$

SCALER 

$$Y_{11} = X_{11}k_{11} + X_{12}k_{12} + X_{21}k_{21} + X_{22}k_{22} + \beta_1$$

$$Y_{12} = X_{12}k_{11} + X_{13}X_{12} + X_{22}k_{21} + X_{23}k_{22} + \beta_1$$

$$Y_{21} = X_{21}k_{11} + X_{22}k_{12} + X_{31}k_{21} + X_{32}k_{22} + \beta_1$$

$$Y_{22} = X_{22}k_{11} + X_{23}k_{12} + X_{32}k_{21} + X_{33}k_{22} + \beta_1$$

$$\frac{\partial L}{\partial \beta_1} = \frac{\partial L}{\partial Y_{11}} \cdot \frac{\partial Y_{11}}{\partial \beta_1} + \frac{\partial L}{\partial Y_{12}} \cdot \frac{\partial Y_{12}}{\partial \beta_1} + \frac{\partial L}{\partial Y_{21}} \cdot \frac{\partial Y_{21}}{\partial \beta_1} + \frac{\partial L}{\partial Y_{22}} \cdot \frac{\partial Y_{22}}{\partial \beta_1}$$

$$= \frac{\partial L}{\partial Y_{11}} + \frac{\partial L}{\partial Y_{12}} + \frac{\partial L}{\partial Y_{21}} + \frac{\partial L}{\partial Y_{22}} \Rightarrow \text{sum}\left(\frac{\partial L}{\partial Y_i}\right)$$

$$\Rightarrow \frac{\partial L}{\partial \beta_1} = \text{sum}\left(\frac{\partial L}{\partial Y_i}\right)$$

Similarly,

$$\frac{\partial L}{\partial \beta_2} = \text{sum}\left(\frac{\partial L}{\partial Y_j}\right)$$

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

$$\Rightarrow \frac{\partial L}{\partial x_{ij}} = \sum \frac{\partial L}{\partial y_{mn}} \cdot \frac{\partial y_{mn}}{\partial x_{ij}}$$

$$y_{11} = x_{11} k_{11} + x_{12} k_{12} + x_{21} k_{21} + x_{22} k_{22} + b_1$$

$$y_{12} = x_{12} k_{11} + x_{13} k_{12} + x_{22} k_{21} + x_{23} k_{22} + b_1$$

$$y_{21} = x_{21} k_{11} + x_{22} k_{12} + x_{31} k_{21} + x_{32} k_{22} + b_1$$

$$y_{22} = x_{22} k_{11} + x_{23} k_{12} + x_{32} k_{21} + x_{33} k_{22} + b_1$$

$$\frac{\partial L}{\partial x^{11}} = \frac{\partial L}{\partial y^{11}} \cdot \frac{\partial y^{11}}{\partial x^{11}}$$

$$\frac{\partial L}{\partial x_{12}} = \frac{\partial L}{\partial y^{11}} \cdot \frac{\partial y^{11}}{\partial x_{12}} + \frac{\partial L}{\partial y^{12}} \cdot \frac{\partial y^{12}}{\partial x_{12}}$$

$$\frac{\partial L}{\partial x_{13}} = \frac{\partial L}{\partial y^{12}} \cdot \frac{\partial y^{12}}{\partial x_{13}}$$

...

$$\begin{aligned} \frac{\partial L}{\partial x_{22}} = & \frac{\partial L}{\partial y^{11}} \cdot \frac{\partial y^{11}}{\partial x_{22}} + \frac{\partial L}{\partial y^{12}} \cdot \frac{\partial y^{12}}{\partial x_{22}} \\ & + \frac{\partial L}{\partial y^{21}} \cdot \frac{\partial y^{21}}{\partial x_{22}} + \frac{\partial L}{\partial y^{22}} \cdot \frac{\partial y^{22}}{\partial x_{22}} \end{aligned}$$

$$\frac{\partial L}{\partial x_{11}} = \frac{\partial L}{\partial y_{11}} \cdot k_{11} ; \quad \frac{\partial L}{\partial x_{12}} = \frac{\partial L}{\partial y_{11}} \cdot k_{12} + \frac{\partial L}{\partial y_{12}} \cdot k_{11}$$

$$\frac{\partial L}{\partial x_{13}} = \frac{\partial L}{\partial y_{12}} \cdot k_{12} ; \quad \frac{\partial L}{\partial x_{21}} = \frac{\partial L}{\partial y_{11}} \cdot k_{21} + \frac{\partial L}{\partial y_{21}} \cdot k_{11}$$

$$\frac{\partial L}{\partial x_{22}} = \frac{\partial L}{\partial y_{11}} \cdot k_{22} + \frac{\partial L}{\partial y_{12}} \cdot k_{21} + \frac{\partial L}{\partial y_{21}} \cdot k_{12} + \frac{\partial L}{\partial y_{22}} \cdot k_{11}$$

$$\frac{\partial L}{\partial x_{23}} = \frac{\partial L}{\partial y_{12}} \cdot k_{22} + \frac{\partial L}{\partial y_{22}} \cdot k_{12} ; \quad \frac{\partial L}{\partial x_{31}} = \frac{\partial L}{\partial y_{21}} \cdot k_{21}$$

$$\frac{\partial L}{\partial x_{32}} = \frac{\partial L}{\partial y_{21}} \cdot k_{22} + \frac{\partial L}{\partial y_{22}} \cdot k_{21} \Rightarrow \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & \frac{\partial L}{\partial y_{11}} & \frac{\partial L}{\partial y_{12}} & 0 \\ 0 & \frac{\partial L}{\partial y_{21}} & \frac{\partial L}{\partial y_{22}} & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \odot \begin{bmatrix} k_{22} & k_{21} \\ k_{12} & k_{11} \end{bmatrix} = \frac{\partial L}{\partial x}$$

$$\frac{\partial L}{\partial x_{33}} = \frac{\partial L}{\partial y_{22}} \cdot k_{22}$$

$\Rightarrow \frac{\partial L}{\partial x} = \text{conv}(\text{Padded } (\frac{\partial L}{\partial y_i}), 180^\circ \text{ Rotated Kernel } K_i)$

Max Pooling Layer

$$\begin{bmatrix} 1 & 1 & 2 & 4 \\ 5 & 6 & 7 & 8 \\ 3 & 2 & 1 & 0 \\ 1 & 2 & 3 & 4 \end{bmatrix} \quad 4 \times 4$$

max Pool 2×2
with $s=2$
→ fwd prop

$$\begin{bmatrix} 6 & 8 \\ 3 & 4 \end{bmatrix} \rightarrow$$

$$\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & \text{dout} & 0 & \text{dout} \\ \text{dout} & 0 & 0 & 0 \\ 0 & 0 & 0 & \text{dout} \end{bmatrix} \quad 4 \times 4$$

← bwd prop $\begin{bmatrix} 6 & 8 \\ 3 & 4 \end{bmatrix}$ ← dout