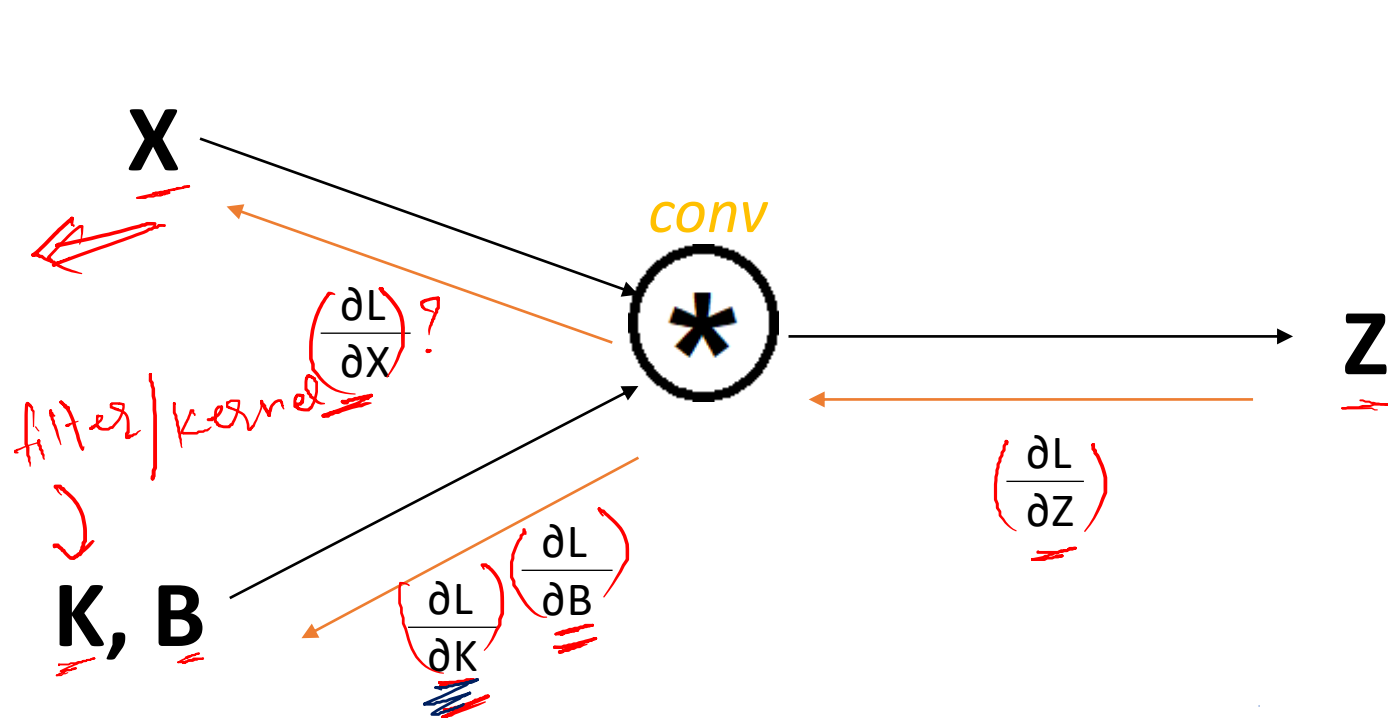


Backpropagation in CNN – Part 1



$$\begin{aligned}
 W &= W - \alpha \frac{\partial L}{\partial W} \\
 K &= K - \alpha \cdot \frac{\partial L}{\partial K} \\
 B &= B - \alpha \cdot \frac{\partial L}{\partial B}
 \end{aligned}
 \quad \left. \vphantom{\begin{aligned} W \\ K \\ B \end{aligned}} \right\}$$

$\rightarrow ()$

$$\frac{\partial L}{\partial K} = \left(\frac{\partial L}{\partial Z} \right) \left(\frac{\partial Z}{\partial K} \right)$$

$$\begin{bmatrix} X_{11} & X_{12} & X_{13} \\ X_{21} & X_{22} & X_{23} \\ X_{31} & X_{32} & X_{33} \end{bmatrix} \odot \begin{bmatrix} \underline{K_{11}} & \underline{K_{12}} \\ K_{21} & \underline{K_{22}} \end{bmatrix} + \underline{B} = \begin{bmatrix} Z_{11} & Z_{12} \\ Z_{21} & Z_{22} \end{bmatrix}$$

$$\begin{aligned}
 \underline{Z_{11}} &= X_{11} \underline{K_{11}} + X_{12} K_{12} + X_{21} K_{21} + X_{22} K_{22} + B \\
 Z_{12} &= X_{12} K_{11} + X_{13} K_{12} + X_{22} K_{21} + X_{23} K_{22} + B \\
 Z_{21} &= X_{21} K_{11} + X_{22} K_{12} + X_{31} K_{21} + X_{32} K_{22} + B \\
 Z_{22} &= X_{22} K_{11} + X_{23} K_{12} + X_{32} K_{21} + X_{33} K_{22} + B
 \end{aligned}$$

$$\frac{\partial L}{\partial K} = \frac{\partial L}{\partial Z} * \frac{\partial Z}{\partial K}$$

$$\frac{\partial L}{\partial K_{mn}} = \sum \left(\frac{\partial L}{\partial Z_{ij}} * \frac{\partial Z_{ij}}{\partial K_{mn}} \right)$$

Einstein Notation

$$\begin{cases} Z_{ij} \rightarrow Z_{11}, Z_{12}, Z_{21}, Z_{22} \\ K_{mn} \rightarrow K_{11}, K_{12}, K_{21}, K_{22} \end{cases}$$

$$\frac{\partial L}{\partial K_{mn}} = \sum \frac{\partial L}{\partial Z_{ij}} * \frac{\partial Z_{ij}}{\partial K_{mn}}$$

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

→ x2

$$\begin{aligned} Z_{11} &= X_{11}K_{11} + X_{12}K_{12} + X_{21}K_{21} + X_{22}K_{22} + B \\ Z_{12} &= X_{12}K_{11} + X_{13}K_{12} + X_{22}K_{21} + X_{23}K_{22} + B \\ Z_{21} &= X_{21}K_{11} + X_{22}K_{12} + X_{31}K_{21} + X_{32}K_{22} + B \\ Z_{22} &= X_{22}K_{11} + X_{23}K_{12} + X_{32}K_{21} + X_{33}K_{22} + B \end{aligned}$$

$$\begin{aligned} \frac{\partial L}{\partial K_{11}} &= \left(\frac{\partial L}{\partial Z_{11}} * \frac{\partial Z_{11}}{\partial K_{11}} \right) + \left(\frac{\partial L}{\partial Z_{12}} * \frac{\partial Z_{12}}{\partial K_{11}} \right) + \left(\frac{\partial L}{\partial Z_{21}} * \frac{\partial Z_{21}}{\partial K_{11}} \right) + \left(\frac{\partial L}{\partial Z_{22}} * \frac{\partial Z_{22}}{\partial K_{11}} \right) \\ \frac{\partial L}{\partial K_{12}} &= \left(\frac{\partial L}{\partial Z_{11}} * \frac{\partial Z_{11}}{\partial K_{12}} \right) + \left(\frac{\partial L}{\partial Z_{12}} * \frac{\partial Z_{12}}{\partial K_{12}} \right) + \left(\frac{\partial L}{\partial Z_{21}} * \frac{\partial Z_{21}}{\partial K_{12}} \right) + \left(\frac{\partial L}{\partial Z_{22}} * \frac{\partial Z_{22}}{\partial K_{12}} \right) \\ \frac{\partial L}{\partial K_{21}} &= \frac{\partial L}{\partial Z_{11}} * \frac{\partial Z_{11}}{\partial K_{21}} + \frac{\partial L}{\partial Z_{12}} * \frac{\partial Z_{12}}{\partial K_{21}} + \frac{\partial L}{\partial Z_{21}} * \frac{\partial Z_{21}}{\partial K_{21}} + \frac{\partial L}{\partial Z_{22}} * \frac{\partial Z_{22}}{\partial K_{21}} \\ \frac{\partial L}{\partial K_{22}} &= \frac{\partial L}{\partial Z_{11}} * \frac{\partial Z_{11}}{\partial K_{22}} + \frac{\partial L}{\partial Z_{12}} * \frac{\partial Z_{12}}{\partial K_{22}} + \frac{\partial L}{\partial Z_{21}} * \frac{\partial Z_{21}}{\partial K_{22}} + \frac{\partial L}{\partial Z_{22}} * \frac{\partial Z_{22}}{\partial K_{22}} \end{aligned}$$

→ x33

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

$$\frac{\partial L}{\partial K_{11}} = \frac{\partial L}{\partial Z_{11}} * X_{11} + \frac{\partial L}{\partial Z_{12}} * X_{12} + \frac{\partial L}{\partial Z_{21}} * X_{21} + \frac{\partial L}{\partial Z_{22}} * X_{22}$$

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

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

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

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

$$\frac{\partial L}{\partial K} = \text{conv}(X, \frac{\partial L}{\partial Z})$$

$$\frac{\partial L}{\partial B} = \frac{\partial L}{\partial z_{ij}} \times \frac{\partial z_{ij}}{\partial B}$$

$$\frac{\partial L}{\partial B} = \sum \frac{\partial L}{\partial z_{ij}} * \frac{\partial z_{ij}}{\partial B}$$

$$= \frac{\partial L}{\partial z_{11}} * \left(\frac{\partial z_{11}}{\partial B} \right) + \frac{\partial L}{\partial z_{12}} * \left(\frac{\partial z_{12}}{\partial B} \right) + \frac{\partial L}{\partial z_{21}} * \left(\frac{\partial z_{21}}{\partial B} \right) + \frac{\partial L}{\partial z_{22}} * \left(\frac{\partial z_{22}}{\partial B} \right)$$

$$\frac{\partial L}{\partial B} = \frac{\partial L}{\partial z_{11}} + \frac{\partial L}{\partial z_{12}} + \frac{\partial L}{\partial z_{21}} + \frac{\partial L}{\partial z_{22}}$$

$$\cancel{\frac{\partial}{\partial B}} z_{11} = \cancel{X_{11} K_{11}} + \cancel{X_{12} K_{12}} + \cancel{X_{21} K_{21}} + \cancel{X_{22} K_{22}} + \underline{B}$$

$$\underline{z_{12}} = X_{12} K_{11} + X_{13} K_{12} + X_{22} K_{21} + X_{23} K_{22} + \underline{B}$$

$$\underline{z_{21}} = X_{21} K_{11} + X_{22} K_{12} + X_{31} K_{21} + X_{32} K_{22} + \underline{B}$$

$$\underline{z_{22}} = X_{22} K_{11} + X_{23} K_{12} + X_{32} K_{21} + X_{33} K_{22} + \underline{B}$$

$$\boxed{\frac{\partial L}{\partial B} = \text{sum} \left(\frac{\partial L}{\partial Z} \right)}$$

$$\frac{\partial L}{\partial X_{mn}} = \sum_i \frac{\partial L}{\partial Z_{ij}} * \frac{\partial Z_{ij}}{\partial X_{mn}}$$

$$\frac{\partial L}{\partial X_{mn}} = \sum \frac{\partial L}{\partial Z_{ij}} * \frac{\partial Z_{ij}}{\partial X_{mn}}$$

$$Z_{11} = X_{11}K_{11} + X_{12}K_{12} + X_{21}K_{21} + X_{22}K_{22} + B$$

$$Z_{12} = X_{12}K_{11} + X_{13}K_{12} + X_{22}K_{21} + X_{23}K_{22} + B$$

$$Z_{21} = X_{21}K_{11} + X_{22}K_{12} + X_{31}K_{21} + X_{32}K_{22} + B$$

$$Z_{22} = X_{22}K_{11} + X_{23}K_{12} + X_{32}K_{21} + X_{33}K_{22} + B$$

$$\begin{bmatrix} X_{11} & X_{12} & X_{13} \\ X_{21} & X_{22} & X_{23} \\ X_{31} & X_{32} & X_{33} \end{bmatrix} \odot \begin{bmatrix} K_{11} & K_{12} \\ K_{21} & K_{22} \end{bmatrix} + B = \begin{bmatrix} Z_{11} & Z_{12} \\ Z_{21} & Z_{22} \end{bmatrix}$$

$$\frac{\partial L}{\partial X_{11}} = \frac{\partial L}{\partial Z_{11}} * \left(\frac{\partial Z_{11}}{\partial X_{11}} \right) = \frac{\partial L}{\partial Z_{11}} * K_{11}$$

$$\frac{\partial L}{\partial X_{12}} = \left(\frac{\partial L}{\partial Z_{11}} * \frac{\partial Z_{11}}{\partial X_{12}} \right) + \left(\frac{\partial L}{\partial Z_{12}} * \frac{\partial Z_{12}}{\partial X_{12}} \right) = \frac{\partial L}{\partial Z_{12}} * K_{12} + \frac{\partial L}{\partial Z_{12}} * K_{11}$$

$$\frac{\partial L}{\partial X_{13}} = \frac{\partial L}{\partial Z_{12}} * \frac{\partial Z_{12}}{\partial X_{13}} = \frac{\partial L}{\partial Z_{12}} * K_{12}$$

Take a pause and note here that,

$$\frac{\partial Z_{12}}{\partial X_{11}}, \frac{\partial Z_{21}}{\partial X_{11}}, \frac{\partial Z_{22}}{\partial X_{11}} = 0$$

as any change in X_{11} does not affect Z_{12}, Z_{21}, Z_{22} .

And thus, they are not considered in this equation

$$\frac{\partial L}{\partial X_{22}} = \left(\frac{\partial L}{\partial Z_{11}} * K_{22} \right) + \left(\frac{\partial L}{\partial Z_{12}} * K_{21} \right) + \left(\frac{\partial L}{\partial Z_{21}} * K_{12} \right) + \left(\frac{\partial L}{\partial Z_{22}} * K_{11} \right) \quad \dots \quad \frac{\partial L}{\partial X_{mn}} \quad \left\{ \begin{matrix} mn \rightarrow 11 \rightarrow 33 \end{matrix} \right.$$

$$\frac{\partial L}{\partial X_{11}} = \frac{\partial L}{\partial Z_{11}} * K_{11}$$

$$\frac{\partial L}{\partial X_{12}} = \frac{\partial L}{\partial Z_{11}} * K_{12} + \frac{\partial L}{\partial Z_{12}} * K_{11}$$

$$\frac{\partial L}{\partial X_{13}} = \frac{\partial L}{\partial Z_{12}} * K_{12}$$

$$\frac{\partial L}{\partial X_{21}} = \frac{\partial L}{\partial Z_{11}} * K_{21} + \frac{\partial L}{\partial Z_{21}} * K_{11}$$

$$\frac{\partial L}{\partial X_{22}} = \frac{\partial L}{\partial Z_{11}} * K_{22} + \frac{\partial L}{\partial Z_{12}} * K_{21} + \frac{\partial L}{\partial Z_{21}} * K_{12} + \frac{\partial L}{\partial Z_{22}} * K_{11}$$

$$\frac{\partial L}{\partial X_{23}} = \frac{\partial L}{\partial Z_{12}} * K_{22} + \frac{\partial L}{\partial Z_{22}} * K_{12}$$

$$\frac{\partial L}{\partial X_{31}} = \frac{\partial L}{\partial Z_{21}} * K_{21}$$

$$\frac{\partial L}{\partial X_{32}} = \frac{\partial L}{\partial Z_{21}} * K_{22} + \frac{\partial L}{\partial Z_{22}} * K_{21}$$

$$\frac{\partial L}{\partial X_{33}} = \frac{\partial L}{\partial Z_{22}} * K_{22}$$

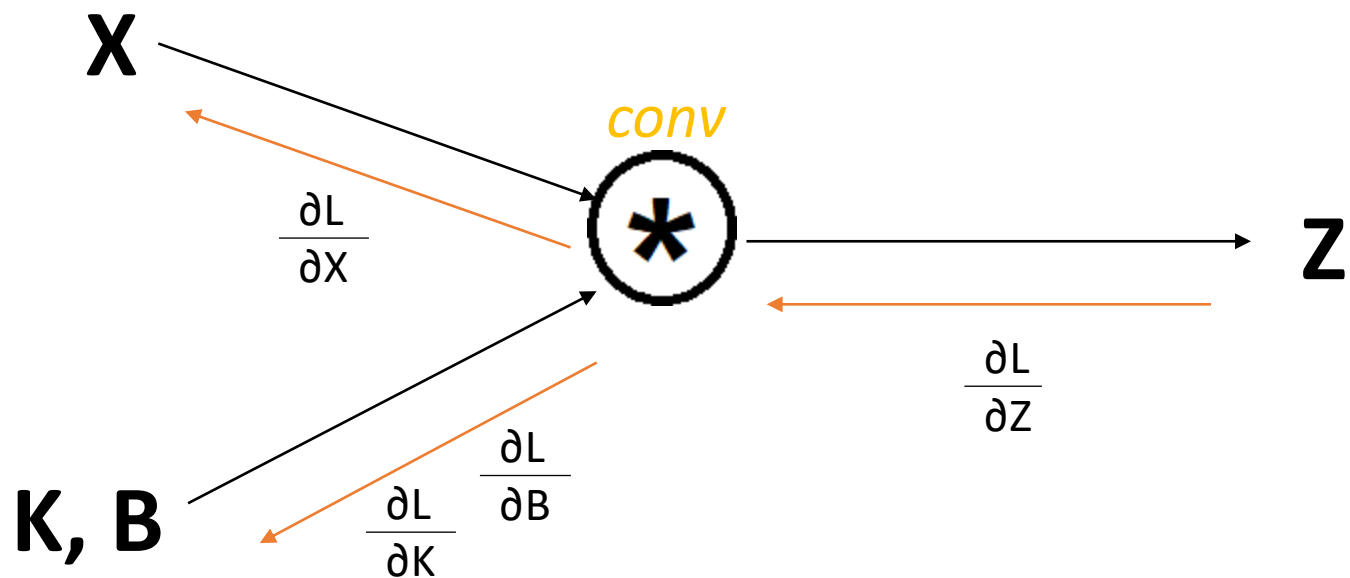
180°

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

$$\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & \frac{\partial L}{\partial Z_{11}} & \frac{\partial L}{\partial Z_{12}} & 0 \\ 0 & \frac{\partial L}{\partial Z_{21}} & \frac{\partial L}{\partial Z_{22}} & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\frac{\partial L}{\partial X} = \text{conv}(\text{padded}(\frac{\partial L}{\partial Z}), \text{180}^\circ \text{ rotated filter } K)$$

$$\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & \frac{\partial L}{\partial Z_{11}} & \frac{\partial L}{\partial Z_{12}} & 0 \\ 0 & \frac{\partial L}{\partial Z_{21}} & \frac{\partial L}{\partial Z_{22}} & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \otimes \begin{bmatrix} K_{22} & K_{21} \\ K_{12} & K_{11} \end{bmatrix} = \begin{bmatrix} \frac{\partial L}{\partial X_{11}} & \frac{\partial L}{\partial X_{12}} & \frac{\partial L}{\partial X_{13}} \\ \frac{\partial L}{\partial X_{21}} & \frac{\partial L}{\partial X_{22}} & \frac{\partial L}{\partial X_{23}} \\ \frac{\partial L}{\partial X_{31}} & \frac{\partial L}{\partial X_{32}} & \frac{\partial L}{\partial X_{33}} \end{bmatrix}$$



$$\frac{\partial L}{\partial K} = \text{conv}(X, \frac{\partial L}{\partial Z})$$

$$K = K - \alpha * \frac{\partial L}{\partial K}$$

$$\frac{\partial L}{\partial B} = \text{sum}(\frac{\partial L}{\partial Z})$$

$$B = B - \alpha * \frac{\partial L}{\partial B}$$

$$\frac{\partial L}{\partial X} = \text{conv}(\text{padded}(\frac{\partial L}{\partial Z}), \text{180}^\circ \text{ rotated filter K})$$