Backdrop through convolution

Tuesday, February 6, 2024 11:41 PM

$$\begin{bmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \end{bmatrix} * \begin{bmatrix} w_{11} & w_{12} \\ w_{21} & w_{22} \end{bmatrix} = \begin{bmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \end{bmatrix}$$

$$y_{11} = w_{11} \cdot x_{11} + w_{12} \cdot x_{12} + w_{21} \cdot x_{21} + w_{22} \cdot x_{22}$$

$$y_{12} = w_{11} \cdot x_{12} + w_{12} \cdot x_{13} + w_{21} \cdot x_{22} + w_{22} \cdot x_{23}$$

$$y_{21} = w_{11} \cdot x_{21} + w_{12} \cdot x_{22} + w_{21} \cdot x_{31} + w_{22} \cdot x_{32}$$

$$y_{22} = w_{11} \cdot x_{22} + w_{12} \cdot x_{23} + w_{21} \cdot x_{32} + w_{22} \cdot x_{33}$$

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

$$\begin{bmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \end{bmatrix}$$

$$\frac{\partial L}{\partial w_{12}} = \frac{\partial L}{\partial y_{11}} \cdot \underbrace{\frac{\partial y_{11}}{\partial w_{12}}}_{x_{12}} + \frac{\partial L}{\partial y_{12}} \cdot \underbrace{\frac{\partial y_{12}}{\partial w_{12}}}_{x_{13}} + \frac{\partial L}{\partial y_{21}} \cdot \underbrace{\frac{\partial y_{21}}{\partial w_{12}}}_{x_{22}} + \frac{\partial L}{\partial y_{22}} \cdot \underbrace{\frac{\partial y_{22}}{\partial w_{12}}}_{x_{23}}$$

$$\begin{bmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \end{bmatrix}$$

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

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

$$y_{1} = w_{11} \cdot x_{11} + w_{12} \cdot x_{12} + w_{21} \cdot x_{21} + w_{22} \cdot x_{22}$$

$$y_{12} = w_{11} \cdot x_{12} + w_{12} \cdot x_{13} + w_{21} \cdot x_{22} + w_{22} \cdot x_{23}$$

$$y_{21} = w_{11} \cdot x_{21} + w_{12} \cdot x_{22} + w_{21} \cdot x_{31} + w_{22} \cdot x_{32}$$

$$y_{22} = w_{11} \cdot x_{22} + w_{12} \cdot x_{23} + w_{21} \cdot x_{32} + w_{22} \cdot x_{3}$$

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

$$= \quad w_{11} \qquad 0 \qquad 0 \qquad 0$$

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

$$= \quad w_{12} \qquad w_{11} \qquad 0 \qquad 0$$

$$\frac{\partial L}{\partial x_{13}} = \quad 0 \qquad w_{12} \qquad 0 \qquad 0$$

$$\frac{\partial L}{\partial x_{21}} = \quad w_{21} \qquad 0 \qquad w_{11} \qquad 0$$

$$\frac{\partial L}{\partial x_{22}} = \quad w_{22} \qquad w_{21} \qquad w_{12} \qquad w_{11}$$

$$\frac{\partial L}{\partial x_{23}} = \quad 0 \qquad w_{22} \qquad 0 \qquad w_{12}$$

$$\frac{\partial L}{\partial x_{31}} = \quad 0 \qquad 0 \qquad w_{21} \qquad 0$$

$$\frac{\partial L}{\partial x_{32}} = \quad 0 \qquad 0 \qquad w_{22} \qquad w_{21}$$

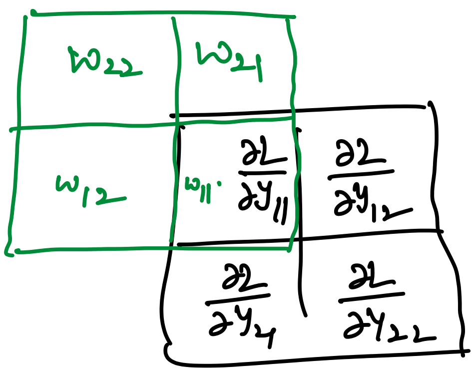
$$\frac{\partial L}{\partial x_{33}} = \qquad\qquad\qquad\qquad w_{22}$$

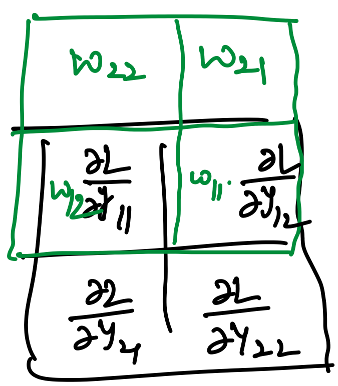
→ Take the 180° rotated version of the filter | weights

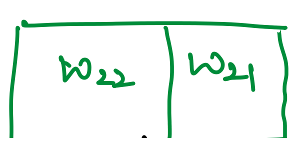
$$\begin{bmatrix} w_{11} & w_{12} \\ w_{21} & w_{22} \end{bmatrix} \longrightarrow \begin{bmatrix} w_{22} & w_{21} \\ w_{12} & w_{11} \end{bmatrix}$$

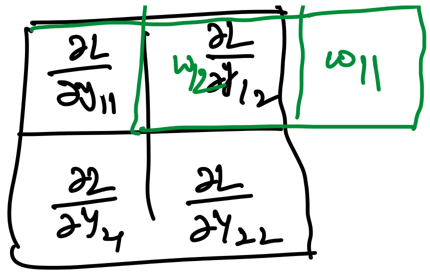
equivalent to taking a vertical flip then a horizontal flip

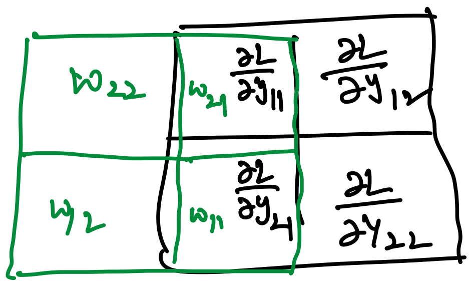
→ Now take a 'full' convolution of the loss gradients w.r.t this flipped filter











one may proceed all the way to the bottom right