

CV大作业理论部分

卷积层:

$$\begin{bmatrix} I_{11} & I_{12} & I_{13} \\ I_{21} & I_{22} & I_{23} \\ I_{31} & I_{32} & I_{33} \end{bmatrix} * \begin{bmatrix} W_{11} & W_{12} \\ W_{21} & W_{22} \end{bmatrix} + b = \begin{bmatrix} O_{11} & O_{12} \\ O_{21} & O_{22} \end{bmatrix}$$

前向传播:

$$O_{ij} = \sum_{b=1}^2 \sum_{a=1}^2 I_{i+a-1, j+b-1} \cdot W_{ab} + b$$

反向传播:

$$(1) \frac{\partial L}{\partial b} = \sum_{i=1}^2 \sum_{j=1}^2 \frac{\partial L}{\partial O_{ij}} \frac{\partial O_{ij}}{\partial b} = \sum_{i=1}^2 \sum_{j=1}^2 \frac{\partial L}{\partial O_{ij}}$$

$$(2) \frac{\partial L}{\partial W_{ij}} = \sum_{a=1}^2 \sum_{b=1}^2 \frac{\partial L}{\partial O_{ab}} \frac{\partial O_{ab}}{\partial W_{ij}} = \sum_{a=1}^2 \sum_{b=1}^2 \frac{\partial L}{\partial O_{ab}} \cdot I_{i+a-1, j+b-1}$$

(3) $\frac{\partial L}{\partial I_{ij}} = \frac{\partial L}{\partial O_{ij}}$ 其实 $\frac{\partial L}{\partial I}$ 的推导实在是不会了!!

计算如下:

(i) 对于左上元素: $I_{11}, I_{13}, I_{31}, I_{33}$

$$\frac{\partial L}{\partial I_{11}} = \frac{\partial L}{\partial O_{11}} \cdot W_{11}, \quad \frac{\partial L}{\partial I_{13}} = \frac{\partial L}{\partial O_{12}} \cdot W_{12}$$

$$\frac{\partial L}{\partial I_{31}} = \frac{\partial L}{\partial O_{21}} \cdot W_{21}, \quad \frac{\partial L}{\partial I_{33}} = \frac{\partial L}{\partial O_{22}} \cdot W_{22}$$

(ii) 对于中心元素: I_{22}

$$\frac{\partial L}{\partial I_{22}} = \frac{\partial L}{\partial O_{11}} \cdot W_{22} + \frac{\partial L}{\partial O_{12}} \cdot W_{21} + \frac{\partial L}{\partial O_{21}} \cdot W_{12} + \frac{\partial L}{\partial O_{22}} \cdot W_{11}$$

(iii) 对于其它元素: $I_{12}, I_{21}, I_{23}, I_{32}$

$$\frac{\partial L}{\partial I_{12}} = \frac{\partial L}{\partial O_{11}} \cdot W_{22} + \frac{\partial L}{\partial O_{12}} \cdot W_{11}$$

$$\frac{\partial L}{\partial I_{21}} = \frac{\partial L}{\partial O_{11}} \cdot W_{21} + \frac{\partial L}{\partial O_{21}} \cdot W_{11}$$

$$\frac{\partial L}{\partial I_{23}} = \frac{\partial L}{\partial O_{12}} \cdot W_{22} + \frac{\partial L}{\partial O_{22}} \cdot W_{12}$$

$$\frac{\partial L}{\partial I_{32}} = \frac{\partial L}{\partial O_{21}} \cdot W_{22} + \frac{\partial L}{\partial O_{22}} \cdot W_{21}$$

池化层 (以最大池化为例)

模型结构与卷积层一致 $\begin{bmatrix} I_{11} & I_{12} & I_{13} \\ I_{21} & I_{22} & I_{23} \\ I_{31} & I_{32} & I_{33} \end{bmatrix} \xrightarrow{\text{pool}} \begin{bmatrix} O_{11} & O_{12} \\ O_{21} & O_{22} \end{bmatrix}$

前向传播:

$$O_{ij} = \max_{\substack{1 \leq a \leq 2 \\ 1 \leq b \leq 2}} \{ I_{i+a-1, j+b-1} \}$$

反向传播:

$$\frac{\partial L}{\partial I} = \frac{\partial L}{\partial O} \cdot \frac{\partial O}{\partial I}$$

(i) 若是取最大值的位置, 则 $\frac{\partial O}{\partial I} = 1$

$$\text{则 } \frac{\partial L}{\partial I} = \frac{\partial L}{\partial O}$$

(ii) 若不是取最大值的位置, 则 $\frac{\partial O}{\partial I} = 0$

$$\text{则 } \frac{\partial L}{\partial I} = 0$$

综上, $\frac{\partial L}{\partial I} = \begin{cases} \frac{\partial L}{\partial O}, & \text{是最大值} \\ 0, & \text{不是最大值} \end{cases}$