

特殊环节的反向传播

①. Relu 函数. $\text{Relu}(x) = \begin{cases} x, & x > 0 \\ 0, & x \leq 0 \end{cases}$ 是连续不可微函数.

Relu 在 $x=0$ 处不可微, 直接将其在 $x=0$ 处的导数置为 1, 则

$$\text{Relu}(x)' = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

最终用 if-else 语句实现反向传播.

②. pooling 池化:

把一个梯度传递给 $n \times n$ 个像素, 需要保证 ^{梯度} 传递的总和不变。

1) mean pooling: 平分 n 份进行传递, 即每个像素 $\frac{1}{n}$

2) max pooling: 前向时记录 pooling 位置, 后向时, 根据位置传递给前一层。

(不应该 是根据输入的大小进行分配吗?)

即最大元素为 1, 否则为 0, 类似于 Relu.

③. 卷积层:

在反向传播过程中, 若第 L 层的 a 节点通过权值 w 对 $L+1$ 层的 b 节点有贡献, 则在反向传播过程中, 梯度通过权值 w 从 b 节点, 传播回 a 节点。都遵循这个规律。

所以, 在卷积层的反向中, 需要找到卷积层 L 中的每个单元和 $L+1$ 层中哪些单元相关。

$$\frac{\partial E}{\partial w_{11}} = \frac{\partial E}{\partial o_{11}} \cdot \frac{\partial o_{11}}{\partial w_{11}} + \frac{\partial E}{\partial o_{12}} \cdot \frac{\partial o_{12}}{\partial w_{11}} + \frac{\partial E}{\partial o_{21}} \cdot \frac{\partial o_{21}}{\partial w_{11}} + \frac{\partial E}{\partial o_{22}} \cdot \frac{\partial o_{22}}{\partial w_{11}}$$

$$o_{11} = w_{11} x_{11} + w_{12} x_{12} + \dots + w_{33} x_{33} + b$$

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

其它 w 类似。(可结合矩阵相乘图)。



