HW2-2

假設：

- (z)：卷積操作的結果
- $a = \sigma(z)$：激活函數的輸出
- y'：最終的迴歸結果
- y：正確的標籤
- $L = (y - y')^2$：均方誤差 (MSE) 損失函數

接著，我們使用Chain rule來推導梯度。

1. 對於 w_{11} 的梯度：

$$\frac{\partial L}{\partial w_{11}} = \frac{\partial L}{\partial y'} \cdot \frac{\partial y'}{\partial a} \cdot \frac{\partial a}{\partial z} \cdot \frac{\partial z}{\partial w_{11}}$$

- $\frac{\partial L}{\partial y'} = 2(y' - y)$（MSE 的梯度）
- $\frac{\partial y'}{\partial a} = w$（假設 $y' = wa$）
- $\frac{\partial a}{\partial z} = \sigma'(z)$（假設 σ 是 sigmoid 函數）
- $\frac{\partial z}{\partial w_{11}} = a_{11}$（根據卷積操作）

將這些部分相乘，我們得到梯度：
$\frac{\partial L}{\partial w_{11}} = 2(y' - y) \cdot w \cdot \sigma'(z) \cdot a_{11}$

2. 對於 a_{00} 的梯度：

$$\frac{\partial L}{\partial a_{00}} = \frac{\partial L}{\partial y'} \cdot \frac{\partial y'}{\partial a} \cdot \frac{\partial a}{\partial z} \cdot \frac{\partial z}{\partial a_{00}}$$

- $\frac{\partial z}{\partial a_{00}} = w_{11} \cdot \sigma'(z) \cdot w_{11}$（根據卷積操作和激活函數的導數）

將這些部分相乘，我們得到梯度：

$$\frac{\partial L}{\partial a_{00}} = 2(y' - y) \cdot w \cdot \sigma'(z) \cdot w_{11} \cdot \sigma'(z) \cdot w_{11}$$