

机器视觉第十一次作业

1、 $n \times m$ 的输入，经过一个 $k \times k$ 的卷积核，步长为 s ，当使用不同的 padding 方法时，输出的特征矩阵维度是多少？

当 $n \times m$ 的输入时，经过 $k \times k$ 的卷积核，步长为 s ，则其输出维度为：

$$n' = \left\lfloor \frac{n - k + 2 \times padding}{s} \right\rfloor + 1$$
$$m' = \left\lfloor \frac{m - k + 2 \times padding}{s} \right\rfloor + 1$$

(中括号表示向下取整)输出通道数则一般由程序员设定。

验证：假设为 48×48 的输入，经过 5×5 的卷积核，步长为 2，padding 为 2，则经过该卷积核得到的输出图像为： $\lfloor (48 - 5 + 2 \times 2) / 2 \rfloor + 1 = 24$ ，则为 24×24 。

对于 48×48 的输入图像，经过 5×5 的卷积核，步长为 2，填充为 2 的卷积操作后，输出图像的尺寸为 24×24 。

📄 ↺ 🍏 🔔

🔄 开启新对话

2、简述深度网络反向传播算法实现过程

反向传播的核心是通过链式法则，从输出层开始，逐层计算损失函数对每一层参数的梯度。在进行前向传播后可以将预测值和真实值进行对比，使用损失函数计算二者之间的误差，随后计算损失函数对输出层权重的梯度，表示如何调整权重减少损失。使用链式法则计算损失函数关于每个权重的梯度，使用梯度下降法更新权重，随后逐层传播误差以及计算权重梯度并更新权重。