

# 池化层是如何进行反向传播的?

## 1. 池化类型

最大池化

平均池化

...

## 2. 池化作用

防止过拟合

突出重要特征

降低模型的参数量

减少网络的计算量

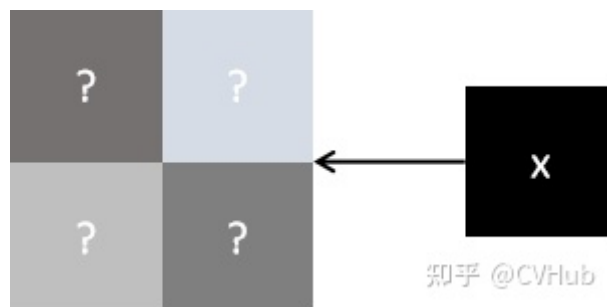
增加非线性表达能力

提升模型的泛化性能

...

## 3. 池化原理

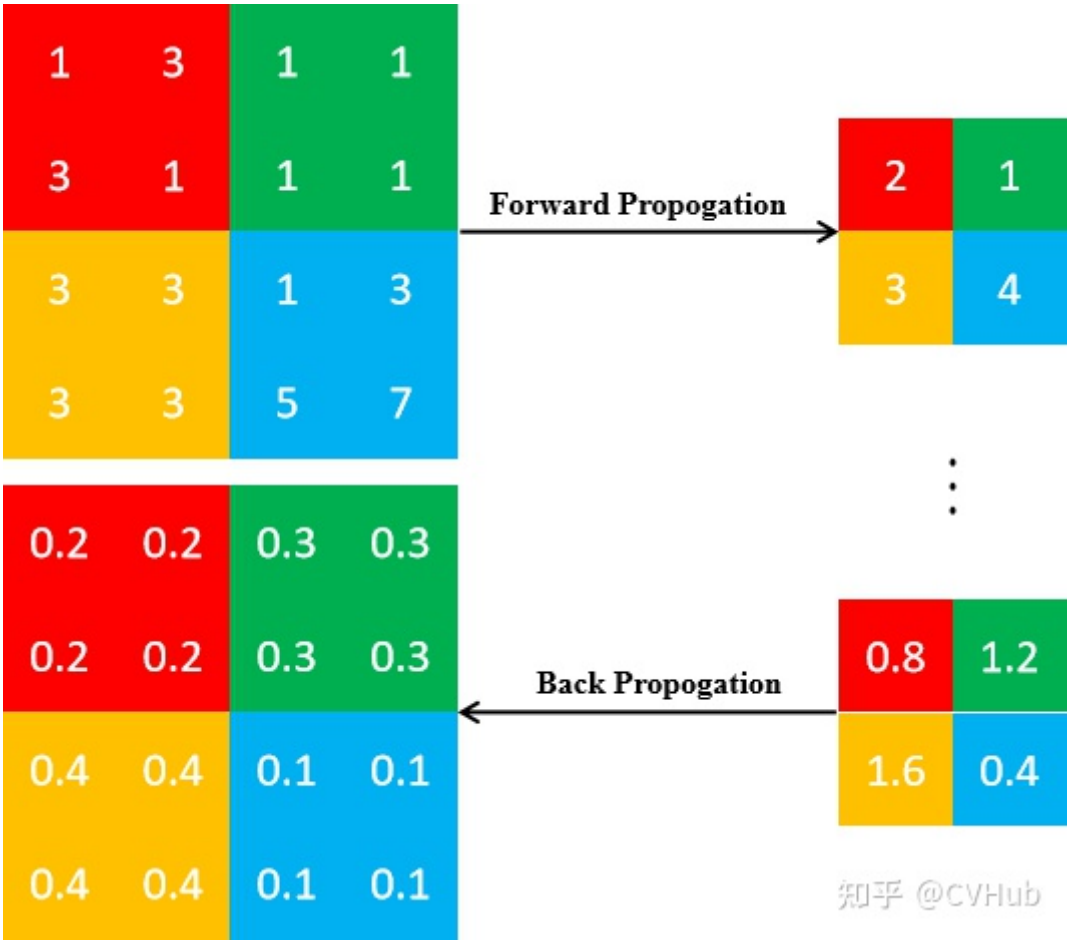
池化操作，通常是用于将特征图进行下采样，减小特征图的尺寸。例如，对于一个  $2 \times 2$  的池化窗口来说，最大池化就是取该窗口内所有元素的最大值，平均池化便是取该窗口内所有元素的平均值。因此，在反向传播的过程中，我们并不能进行逆向操作，即池化层是不可导的。



反向传播过程中，主要涉及到一个问题是梯度回传。梯度通常来说是按照位置索引传播的，所以关键之处是要记录好相应的位置索引，并保证梯度和不变。

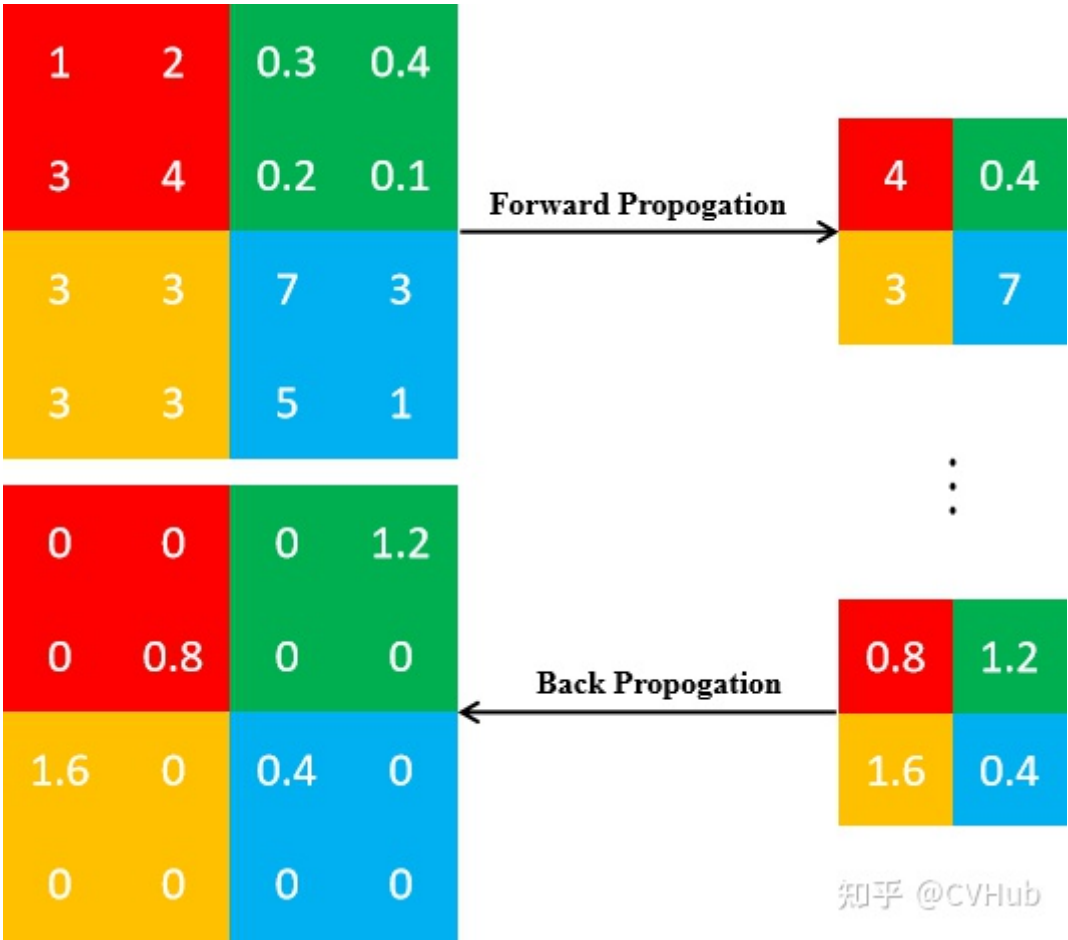
### 3.1 平均池化反向传播原理

对于平均池化来说，由于当前位的元素是由区域内所有元素取平均值所得，因此该梯度和为所有元素的梯度之和。因此，我们在反向传播的过程中只需将当前位的梯度值均分即可。



### 3.2 最大池化反向传播原理

对于最大池化而言，由于当前位的元素是由区域内所有元素取最大值所得，因此该梯度和仅取决于该位的元素，即其余元素的梯度应置为0。因此，我们在前向传播的过程中应该记录该最大值的位置索引。



知乎 @CVHub