

面经 | 计算机视觉面试复习笔记（二）卷积神经网络

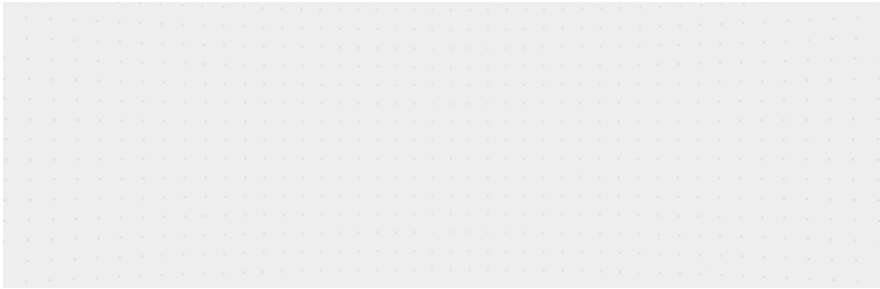
极市平台 2021-10-20 22:00:00 手机阅读 眼

以下文章来源于AI约读社，作者南山



AI约读社
一个热衷于AI学习和分享的平台，期待与您相遇。

↑ 点击[蓝字](#) 关注极市平台



作者 | 南山
来源 | AI约读社
编辑 | 极市平台

极市导读

本文总结了一些关于卷积神经网络的面试知识点。 >>加入极市CV技术交流群，走在计算机视觉的最前沿

计算机视觉面试复习笔记：深度学习基础-神经网络

二、卷积神经网络

在介绍卷积神经网络之前我们先了解一下为什么不使用我们上一节介绍的神经网络（全连接层结构）来处理图像。**使用全连接层结构处理图像时会出现以下问题：**

- 输入图像需要将图像通过 Flatten 层拉成一维，这会导致**丢失空间信息**（像素位置和channel）
- 全连接层考虑的是全局的信息，但在分类图像中，重要的是物体本身的**局部信息**
- 对图像做分类时，**会产生大量的参数，导致计算量过高**。例如：对1000×1000 像素的图片，我们就需要处理3百万个参数。

2.1 卷积神经网络的组成

卷积神经网络通过卷积层提取图像特征，不仅大幅的减少了参数量，还保留了图像的空间信息。

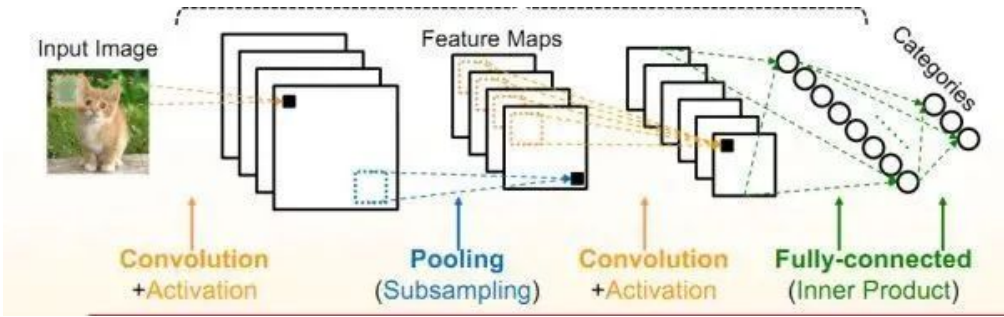
From a few to >100 layers



月发文数目： **
月平均阅读： **

文章工具

- 已发
- 采集图文 合成多
 - 采集样式 查看



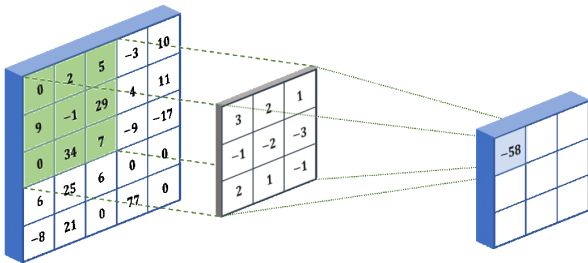
如上图简单的卷积神经网络所示，卷积神经网络主要架构是由卷积层、池化层、全连接层组成。

- 卷积层负责提取图像中的局部特征，卷积层后面一般会加上ReLU层进行非线性激活
- 池化层用来大幅降低参数量级(降维)
- 全连接层类似神经网络的部分，用来输出想要的结果。

2.2 卷积层

卷积层负责提取图像中的局部特征，其原理是通过许多的卷积核(filter, kernel) 在图片上进行滑动提取特征。

下图是卷积滑动的过程，左中右分别为输入层、卷积核、输出层



卷积核 (filter, kernel)

卷积核里面的数字就是卷积层的权重，是经由神经网络训练学习而来的。而卷积核的大小 (kernel size) 及数量 (输出 channel) 是可以调整的超参数，通常会设定为奇数，其原因有两个：

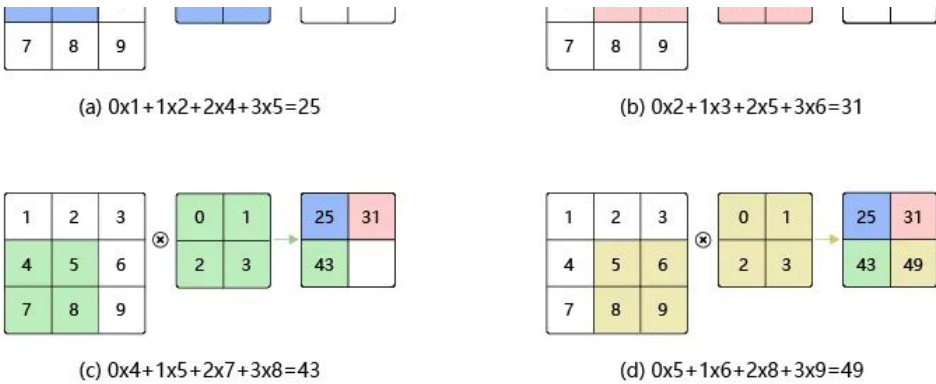
1. 可以保证padding时候，图像的两边依然相对称
2. 在做目标检测时，能够获得中心点，可以更好的预测目标位置

输出层为卷积运算后的结果，称为特征图 (feature map)

卷积运算

卷积运算的方式就是将滑动的窗口与卷积核进行点对点 (elementwise) 的相乘，再将乘完的值相加





卷积的参数

- Kernel size: 定义卷积核的大小，影响卷积操作的感受野，一般使用3x3，5x5 等
- Stride: 定义遍历图像时卷积核移动的步长
- Channel: 定义卷积运算的输入和输出通道数
- Padding: 定义如何处理样本边界的方式，分为不填充或者对边界填充0，不填充的只对特征图做卷积操作，会使得输出的特征图小于输入的特征图；对边界填充0，可以使得输入和输出的特征图保持一致

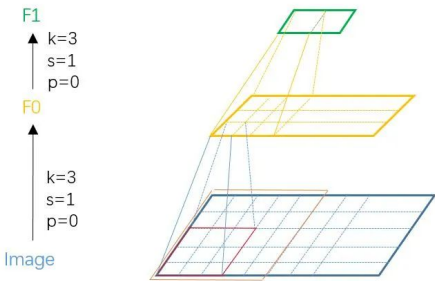
卷积的感受野

感受野是每一层卷积神经网络输出的特征图上的特征点映射到原始图像上的区域的大小，即特征点可以“看到”的范围。描述的原始图像信息，能够表达的信息越全面。

感受野的计算公式如下：

$$l_k = l_{k-1} + \left((f_k - 1) * \prod_{i=1}^{k-1} s_i \right)$$

其中，L表示感受野的大小， l_{k-1} 是第k-1层的感受野大小，而 f_k 是当前层的卷积核大小， s_i 是第i层的步长。**卷积层（F0）**的感受野大小等于其内核k的大小。



F0层： $L_0 = f = 3$

F1层： $L_1 = 3 + (3-1) * 1 = 5$ ；

具体来说，当前特征图的感受野与上层空间有关，与当前层核心大小，与填充和步幅相关。感受野的大小可以大于网络输入的大小。

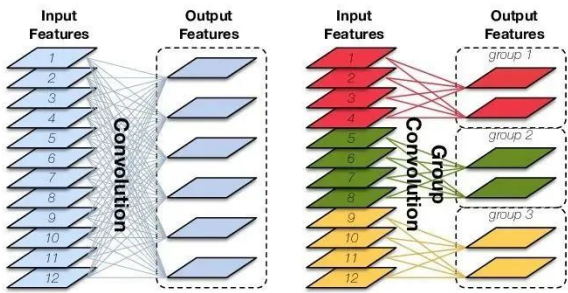
卷积的类型

常见的卷积类型有很多，根据其操作的区域大致可以分为两类：通道相关性、空间相关性。

- 通道相关性的卷积核改变了卷积在channel维度操作，如：Group Convolution、Depthwise Separable Convolutions；
- 空间相关性的卷积核是改变了卷积在w,h维度的操作。除了这两类近年来也出现了一些新的改进思路：如动态卷积，空洞卷积。

Group Convolution（分组卷积）

Group Convolution（分组卷积）就是对输入feature map在channel维度进行分组，然后每组分别卷积。设分成G组，参数量减少为原来的1/G。如下图所示：

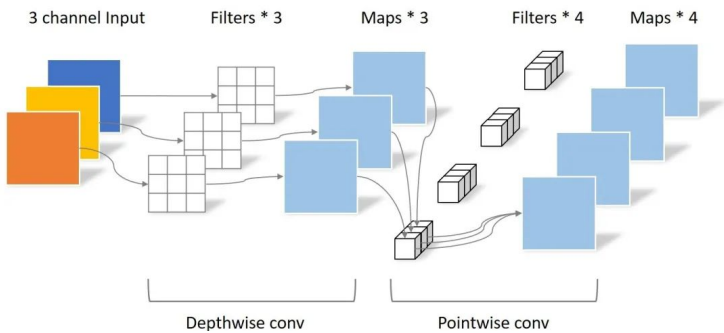


Depthwise Convolution

当分组卷积分组数量等于输入map数量，输出map数量也等于输入map数量，即 $G=C=N$ 、 N 个卷积核每个尺寸为 $1 \times K \times K$ 时，Group Convolution就成了Depthwise Convolution（深度卷积）。

Depthwise Separable Convolution

Depthwise Separable Convolutions（深度可分离卷积）是由 Depthwise Separable Convolution是将一个完整的卷积运算分解为两步进行，即Depthwise Convolution与Pointwise Convolution。



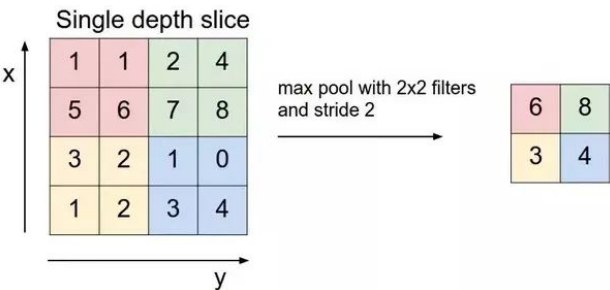
Pointwise Convolution的运算与常规卷积运算非常相似，它的卷积核的尺寸为 $1 \times 1 \times M$ ， M 为上一层的通道数。所以这里的卷积运算会将上一步的map在深度方向上进行加权组合，生成新的Feature map。有几个卷积核就有几个输出Feature map。

2.3 池化层 & 下采样层 pooling

池化层是用来大幅降低参数量(降维)、减少过拟合问题、缓解卷积层对位置的敏感度。空间池化层也被称为子采样或下采样层，它降低了每个特征图的维度但保留了重要信息，简单来说池

化层主要用来缩小特征图的大小减少计算量。池化层主要有以下两种类型：

- 最大池化(max pooling)：就是在框选的局部数值中挑出最大值
- 平均池化（Average pooling）：就是在框选的局部数值做求和去平均

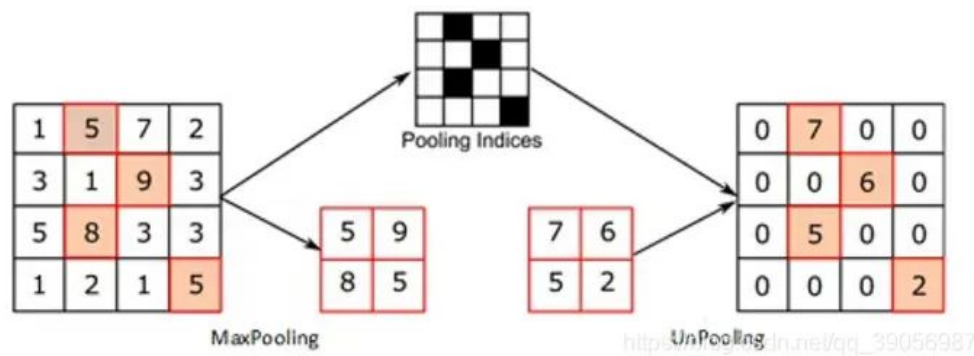


最大池化

上采样

由于输入图像通过卷积神经网络(CNN)提取特征后，输出的尺寸往往会变小，而有时我们需要将图像恢复到原来的尺寸以便进行进一步的计算(如图像的语义分割)，这个使图像由小分辨率映射到大分辨率的操作，叫做上采样，它的实现一般有三种方式：

- 插值，插值法不需要学习任何的参数，只是根据已知的像素点对未知的点进行预测估计，一般使用的是双线性插值，其他插值方式还有最近邻插值、三线性插值等；
- 转置卷积又或是说反卷积，通过对输入feature map间隔填充0，再进行标准的卷积计算，可以使得输出feature map的尺寸比输入更大；
- Max Unpooling，在对称的max pooling位置记录最大值的索引位置，然后在unpooling阶段时将对应的值放置到原先最大值位置，其余位置补0；



2.4 展平层 Flatten Layer

Flatten层的作用就是将卷积层与池化层输出的特征展平、做维度的转换，如此一来才能放入全连接层做分类

2.5 全连接层

全连接层，其作用是用来进行分类。将卷积层与池化层输出的特征输入到全连接层，通过调整权重及偏差得到分类的结果。

如果觉得有用，就请分享到朋友圈吧！



极市平台

为计算机视觉开发者提供全流程算法开发训练平台，以及大咖技术分享、社区交流、竞...
765篇原创内容

公众号

△点击卡片关注极市平台，获取最新CV干货

公众号后台回复“**CVPR21检测**”获取CVPR2021目标检测论文下载~

极市干货

神经网络：视觉神经网络模型优秀开源工作：timm库使用方法和最新代码解读

技术综述：综述：神经网络中 Normalization 的发展历程 | CNN轻量化模型及其设计原则综述

算法技巧（trick）：8点PyTorch提速技巧汇总 | 图像分类算法优化技巧



CV技术社群邀请码



△长按添加极市小助手

添加极市小助手微信（ID：cvmart4）

备注：姓名-学校/公司-研究方向-城市（如：小极-北大-目标检测-深圳）

即可申请加入极市 目标检测/图像分割/工业检测/人脸/医学影像/3D/SLAM/自动驾驶/超分辨率/姿态估计/ReID/GAN/图像增强/OCR/视频理解等技术交流群

每月大咖直播分享、真实项目需求对接、求职内推、算法竞赛、干货资讯汇总、与 10000+来自港科大、北大、清华、中科院、CMU、腾讯、百度等名校名企视觉开发者互动交流~

觉得有用麻烦给个在看啦~ 

喜欢此内容的人还喜欢

我是如何从头开始写一篇顶级论文的

计算机视觉与机器学习



产品渲染练习（前半段）



小斯设计工作室



有趣大会 · ACL2022 (长论文篇)
看个通俗理解吧

