

目录



- ◉ 前馈神经网络
- ◉ 卷积神经网络
 - ◆ 卷积运算
 - ◆ 卷积层、池化层、感受野
 - ◆ 典型的网络结构
 - ◆ 实际应用
- ◉ 循环神经网络
- ◉ 如何实现

7



卷积神经网络



- ◉ 卷积神经网络 (Convolutional Neural Network, CNN)
 - ◆ 应用广泛的一类前馈神经网络
 - ◆ 在计算机视觉诸多任务中取得突破性成果
 - ◆ CNN逐渐在自然语言处理、推荐系统和语音识别等领域得到广泛应用

8



卷积神经网络



核心操作：卷积运算

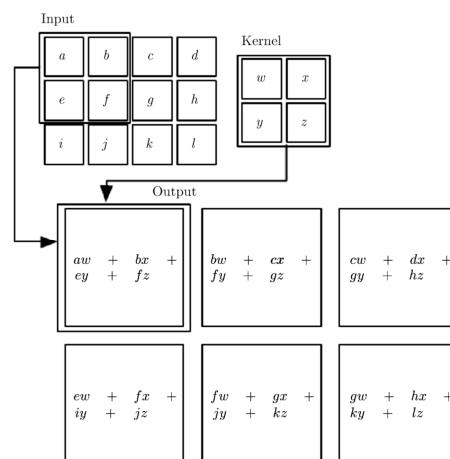
- ◆ 卷积核，很重要！
- ◆ 在输入区域上进行“从左至右、从上至下”的平移操作
- ◆ 每到达一个位置时，计算局部区域的输出值

特点

- ◆ 局部连接
- ◆ 权值共享

优点

- ◆ 参数规模更小。由于不同区域共享权值，因此仅需少量的参数便能实现全域连接；
- ◆ 平移不变性。计算过程以位置平移为牵引，平移不变



9

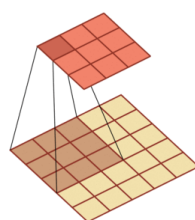


卷积神经网络

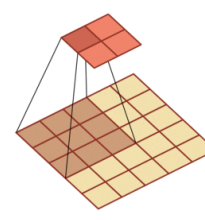


卷积运算中的两个重要超参数

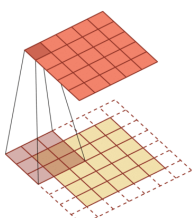
- ◆ 步长 (Stride)
 - 卷积核每做一次运算往前移动的距离
 - 一般设置为1
- ◆ 填充 (Pad)
 - 卷积之前对输入数据周围的填充范围
 - 为什么要填充？
 - 卷积核的边界不能超过输入数据的范围，这使得输出数据的尺寸会略小于输入数据
 - 为了使得输出尺寸与输入保持一致，则需要提前在输入数据的四周填充一定的临时空间
 - 如何填充？“0填充”、“1填充”、“均值填充”等



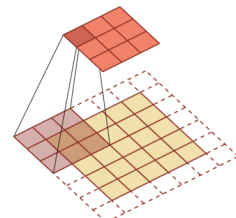
步长1，零填充0



步长2，零填充0



步长1，零填充1



步长2，零填充1

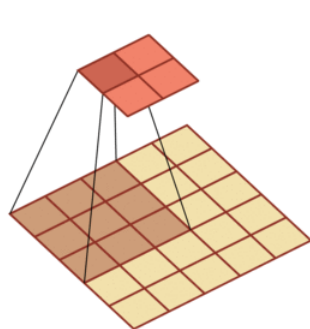
10



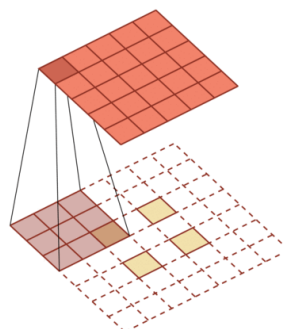
卷积神经网络



- 其他卷积操作 - 反卷积（转置卷积 / 微步卷积）
 - 低维特征映射到高维特征



正卷积



反卷积

11



卷积神经网络-卷积层



- 核心操作在于卷积运算
- 处理不同类型的数据时，需要采用不同的卷积模式
- 根据自由度确定卷积的不同模式
 - 自由度：确定卷积核大小时可以设置的参数数量
 - 自由度为1、2、3：一维、二维和三维卷积
 - n维卷积

12



卷积神经网络-卷积层

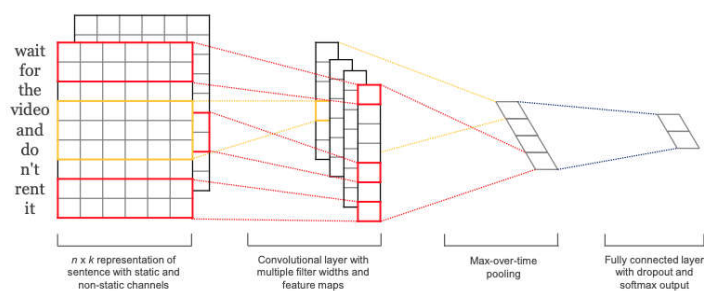


一维卷积

- ◆ 仅用一个参数便可以确定卷积核大小
- ◆ 通常被用于文本等时序数据的学习过程

文本卷积, Text-CNN

- ◆ 每次卷积运算时需要考虑的词窗口大小
- ◆ 提取相邻词之间的关系



13

Figure 1: Model architecture with two channels for an example sentence. dx.doi.org/10.1101/111111

卷积神经网络-卷积层

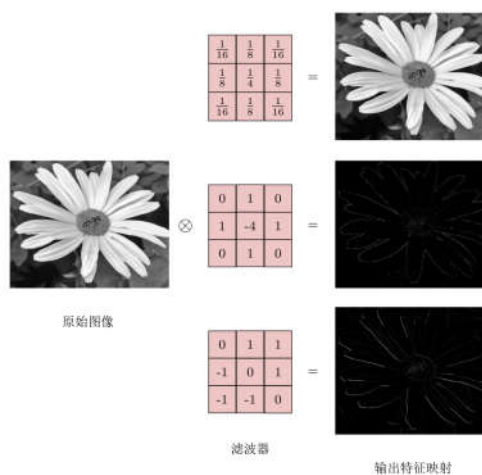


二维卷积

- ◆ 核大小由长度和宽度两个自由参数确定
- ◆ 常被用于图像等二维数据的处理中
- ◆ 目前应用最为广泛的一类卷积模式

图像卷积

- ◆ 不同的卷积核, 多种图像特征



14



卷积神经网络-卷积层

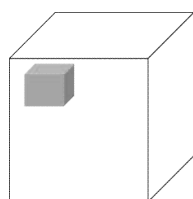


三维卷积

- ◆ 核大小由长度、宽度和深度三个参数确定
- ◆ 常被用于视频等三维数据的处理中

视频卷积

- ◆ 长宽用于模拟视频中每一帧图像的空间维度，深度用于模拟时间维度



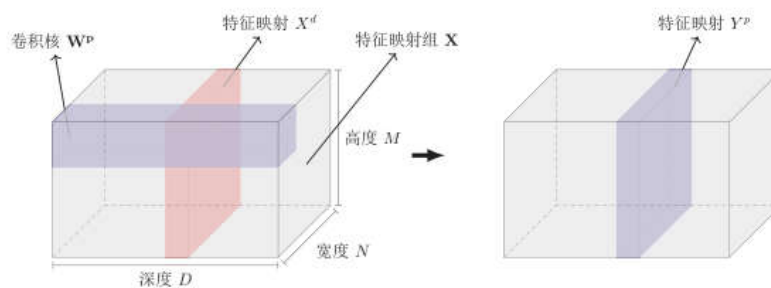
15



卷积神经网络-卷积层



n维卷积



16



卷积神经网络-池化层

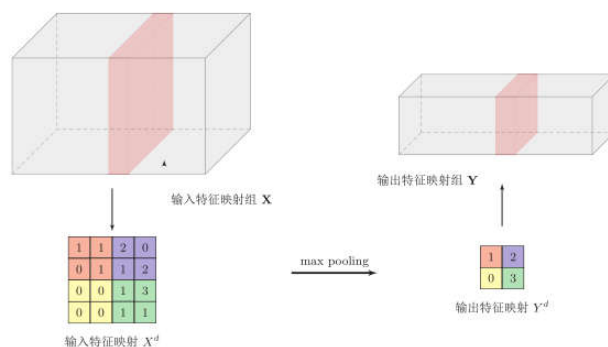


池化层

- ◆ 卷积层并不会显著减少特征映射的神经元个数（即输出数据大小不会显著减少）
- ◆ 将一定区域特征的均值或者最值作为此区域的输出值，简化特征数量
- ◆ 最大池化层 & 最小池化层 & 平均池化层

特点

- ◆ 没有任何参数，不会显著增大性能开销
- ◆ 可以降低数据噪声。去除或平滑图像噪点值
- ◆ 输出单元数量减少，缓解过拟合



17

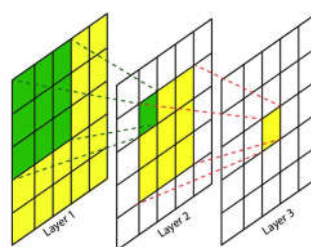
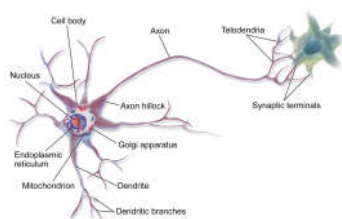


卷积神经网络



感受野

- ◆ 一个感觉神经元的感受野是指这个位置里适当的刺激能够引起该神经元反应的区域
- ◆ 某个输出位置的值受输入数据上某区域的影响，此区域可称为感受野大小
- ◆ 增大感受野的常用方法
 - 增加卷积核大小
 - 增加网络层数
 - 结合使用池化层



18



卷积神经网络

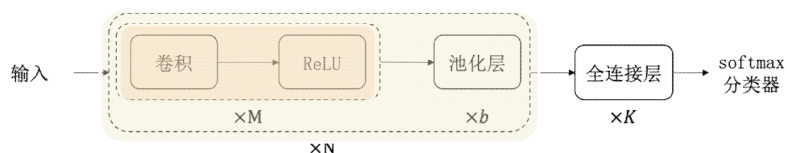


卷积神经网络结构

- ◆ 由卷积层、池化层、全连接层交叉堆叠而成
- ◆ 趋向于小卷积核、大网络深度（一层5*5的卷积 等效于 两层3*3的卷积）

典型结构

- ◆ 一个卷积块主要构成为
 - M 个卷积层（常见设置为2 ~ 5）
 - b个池化层（常见设置为 0或1）
- ◆ 一个卷积网络中可以堆叠
 - N个连续的卷积块（常见设置为1 ~ 100或者更大）
 - 再接着K个全连接层（常见设置为0 ~ 2）



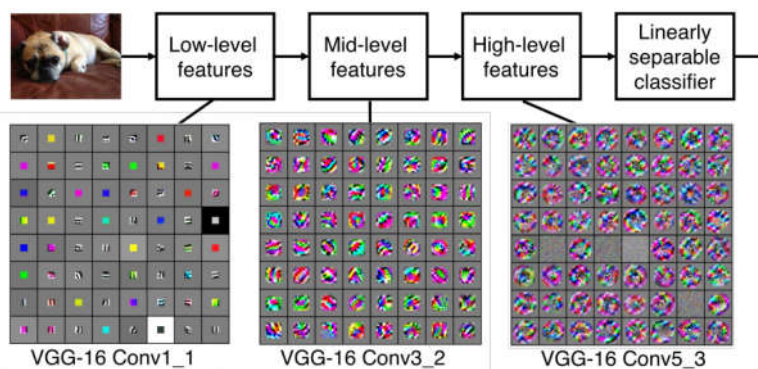
19



卷积神经网络



- 学到了啥
- 浅层：边缘纹理等浅层形状特征，深层：概念化的深层语义特征
- 黑盒特性，难以对每层的结果进行逐一解释



20

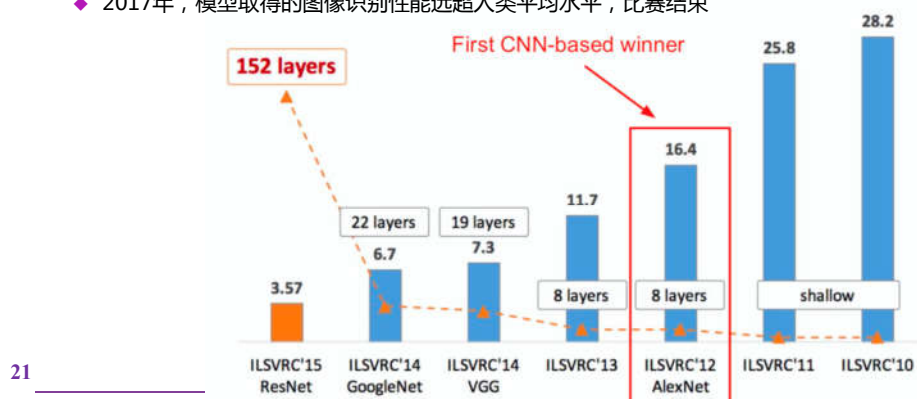


典型的卷积结构



● ILSVRC比赛 (ImageNet竞赛, 大规模视觉识别挑战赛)

- ◆ 自2010年举办以来, 催生了一堆经典的卷积神经网络
- ◆ 2012年, 基于卷积神经网络的模型取得冠军。AlexNet, 8 layers
- ◆ 2014年, ResNet, 深度卷积, 152 layers
- ◆ 2017年, 模型取得的图像识别性能远超人类平均水平, 比赛结束



21

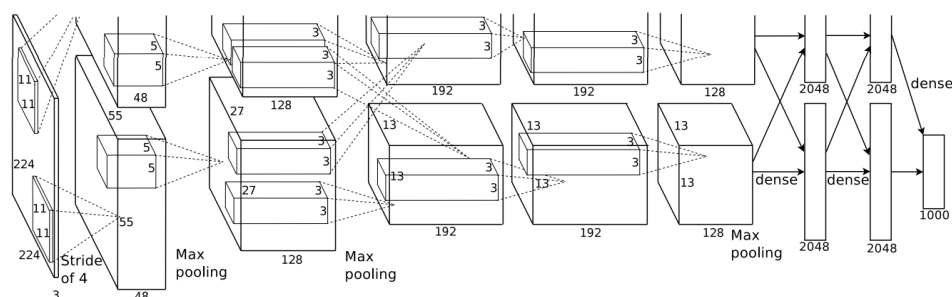


典型的卷积结构



● AlexNet

- ◆ 2012 ILSVRC winner: 第一个基于卷积网络的冠军
- ◆ 共有8层, 其中前5层卷积层, 后边3层全连接层



22

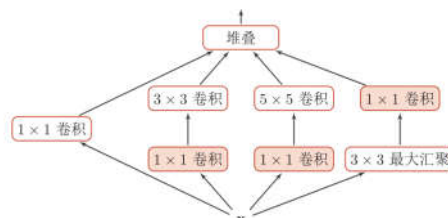


典型的卷积结构

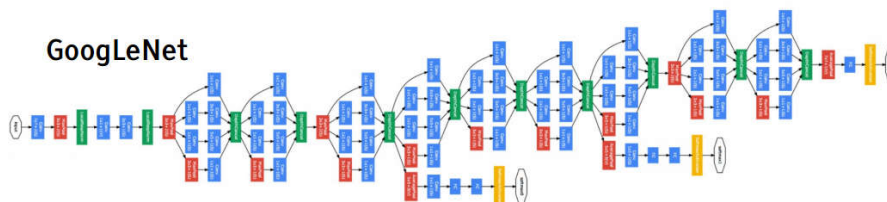


Inception网络

- ◆ 2014 ILSVRC winner
- ◆ 共有22层（更宽 + 更深）



GoogLeNet



23

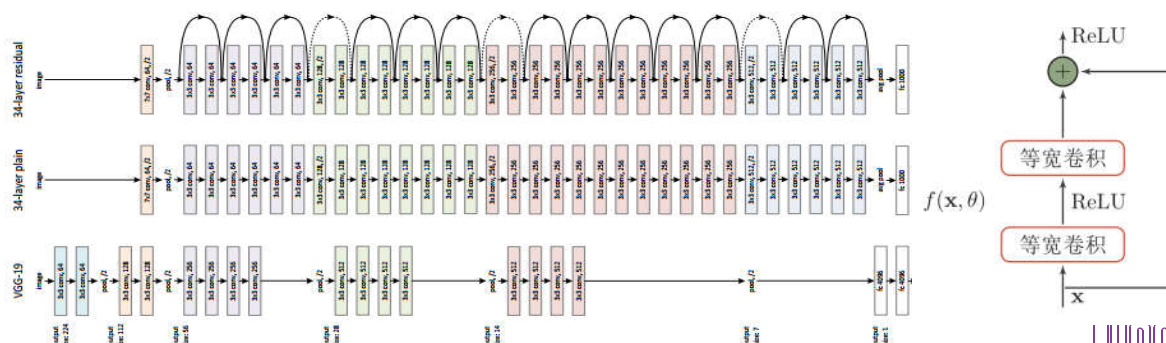


典型的卷积结构



ResNet网络

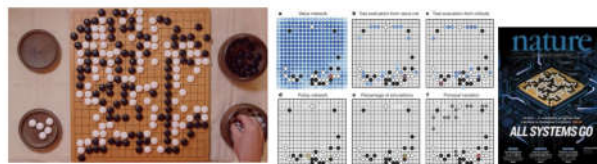
- ◆ 2015 ILSVRC winner
- ◆ 共有152层
- ◆ 深度太深，容易梯度消失/爆炸，网络退化等等
- ◆ 解决方案：残差连接，跳跃连接



卷积神经网络的应用



AlphaGo



The input to the policy network is a $19 \times 19 \times 48$ image stack consisting of 48 feature planes. The first hidden layer zero pads the input into a 23×23 image, then convolves k filters of kernel size 5×5 with stride 1 with the input image and applies a rectifier nonlinearity. Each of the subsequent hidden layers 2 to 12 zero pads the respective previous hidden layer into a 21×21 image, then convolves k filters of kernel size 3×3 with stride 1, again followed by a rectifier nonlinearity. The final layer convolves 1 filter of kernel size 1×1 with stride 1, with a different bias for each position, and applies a softmax function. The match version of AlphaGo used $k = 192$ filters; Fig. 2b and Extended Data Table 3 additionally show the results of training with $k = 128, 256$ and 384 filters.

policy network:

[19x19x48] Input

CONV1: 192 5x5 filters, stride 1, pad 2 => [19x19x192]

CONV2..12: 192 3x3 filters, stride 1, pad 1 => [19x19x192]

CONV: 1 1x1 filter, stride 1, pad 0 => [19x19] (probability map of promising moves)

25



卷积神经网络的应用



实例分割 (Mask RCNN)

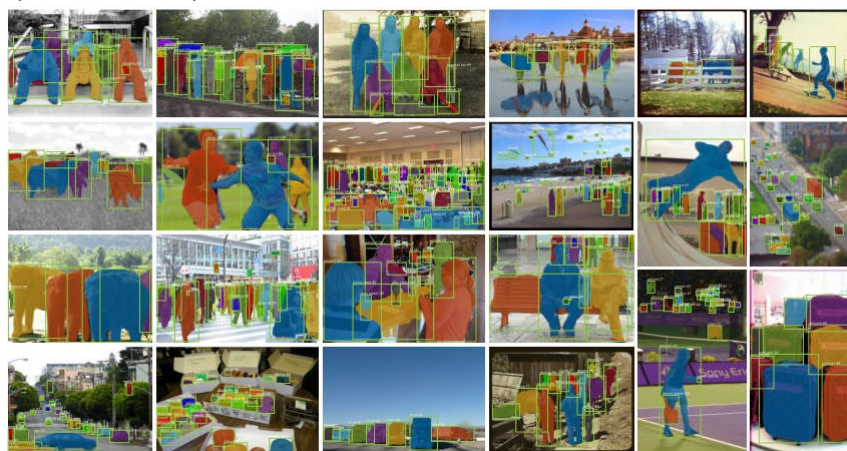


Figure 4. More results of Mask R-CNN on COCO test images, using ResNet-101-FPN and running at 5 fps, with 35.7 mask AP (Table 1).

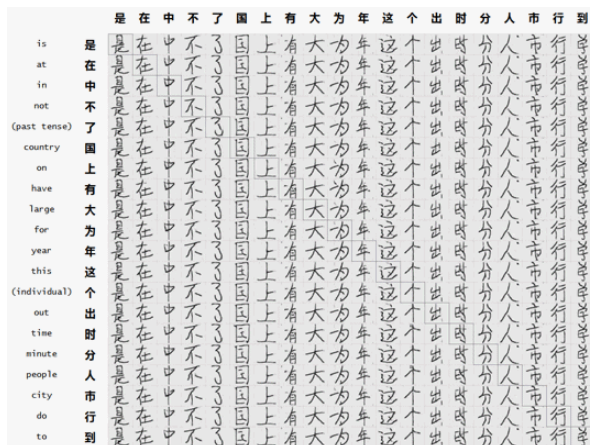
26



卷积神经网络的应用



● 图像生成



27



卷积神经网络的应用



● 画风迁移



28

