

卷积神经网络

赵海臣

背景

- ▶ 卷积神经网络是多层感知器的变种，其来源是对于猫初级视皮层的研究。
- ▶ 初级视皮层包括
 - 简单细胞：主要感知其感受野内的特定边缘刺激
 - 复杂细胞：以简单细胞的输出为输入，以更大的感受野来响应边缘刺激，但忽略刺激的具体位置
- ▶ 卷积神经网络主要采用三种结构来实现对灵长类动物视皮层的模拟
 - 局部连接 / 局部接受域
 - 权值共享
 - 子采样

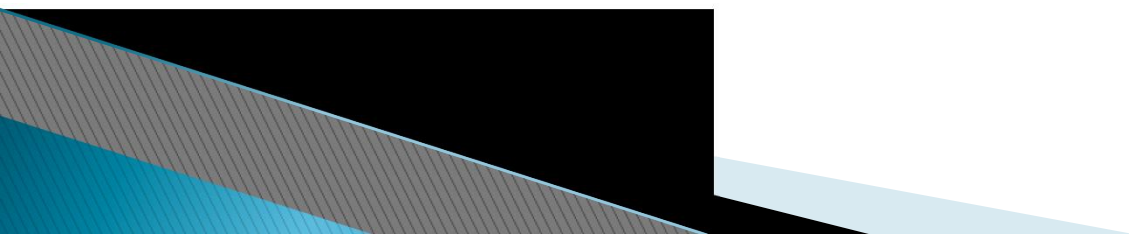
卷积神经网络的特点

▶ 局部连接

- 相邻两层的神元之间的连接采用的是局部连接而不是全连接；

▶ 权值共享

- 在同一层中的部分神经元的权值是共享的



局部连接(即卷积的具体实现)

- 每一层的神经元只与前一层的神经元存在局部连接
 - 例如，第 m 层的神经元只与 $m-1$ 层的神经元的局部区域有连接，第 m 层的感受野的宽度为3，即第 m 层的每个神经元只与第 $m-1$ 层的3个相邻的神经元相连，第 $m+1$ 层与第 m 层的连接也有类似的规则。可以看到 $m+1$ 层的神经元虽然相对于第 m 层的接受域的宽度也为3，但是其相对于第 $m-1$ 层的接受域却为5

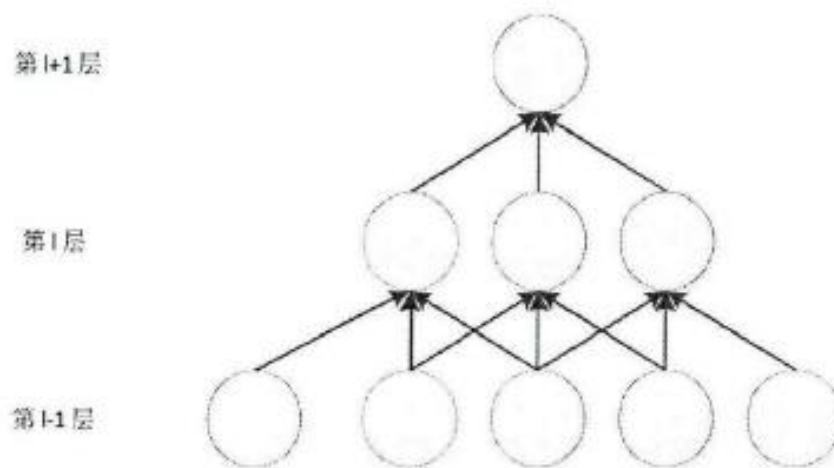
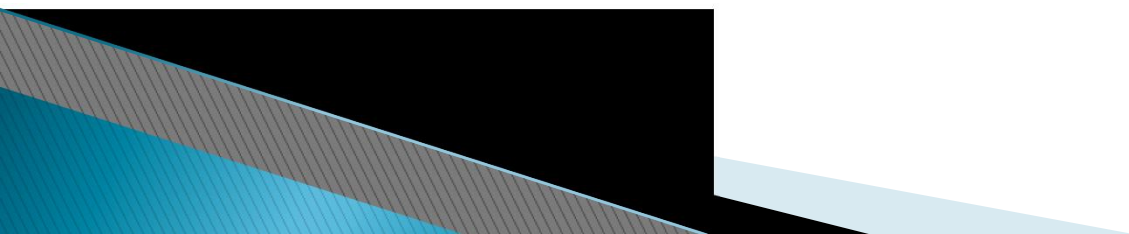


图 2.1 局部连接示意图

Fig. 2.1 Diagram of local connection

局部连接(即卷积的具体实现)

- 这种结构将学习到的过滤器限定在局部空间里(因为每个神经元对其感受野之外的神经元不做反应), 减少了神经元之间的连接数目
- 多个这样的层堆叠在一起之后, 会使得过滤器逐渐成为全局的, 覆盖到更大的区域。



权值共享(即卷积核的具体实现)

- ▶ 权值共享(卷积核的实现)使得共享同一权值的神经元在不同位置检测同一特征，将共享同一权值的神经元组织成一个二维平面，得到特征图。

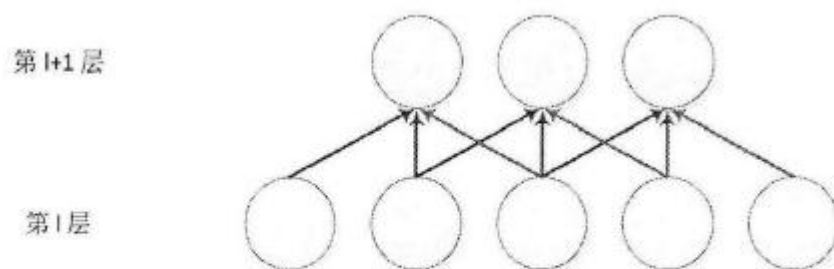


图 2.2 权值共享示意图

Fig. 2.2 Diagram of sharing weights

子采样

- ▶ 子采样操作是在水平和竖直方向以步长为 S 对特征图中的所有 $W \times M$ 大小的连续子区域进行特征映射
- ▶ 一般情况下，映射的过程通常为最大值映射或者是平均值映射，即，在 $W \times M$ 的子区域中，选取最大值或者计算子区域中的平均值作为映射值
 - 如图所示，特征图的大小为 6×6 ，若按以步长为2对特征图中所有大小为 2×2 的连续子区域进行子采样，采样后特征图的大小为 $(6/2) \times (6/2)$ ，即 3×3 。

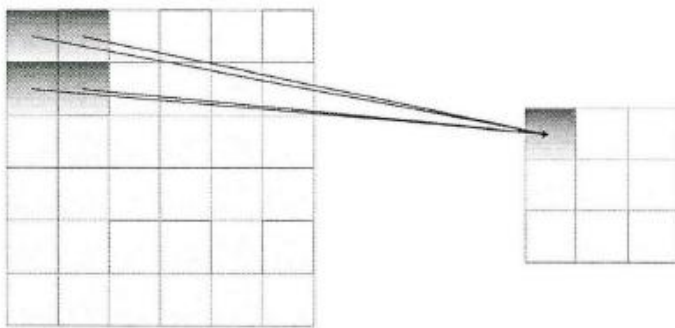


图 2.3 子采样示意图

Fig. 2.3 Diagram of sub sampling

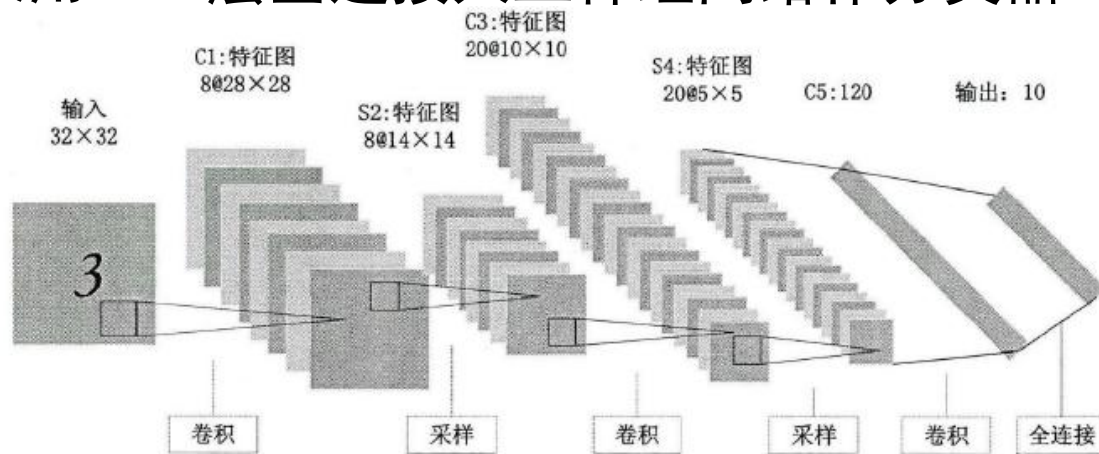
卷积神经网络结构(论文1结构)

▶ 单个特征提取层

- 卷积层
- 子采样层
- 通常1~3个特征提取层

▶ 分类器

- 输入经过逐层提取特征学习到更高层的特征之后，仅最后一个阶段得到的特征被输入分类器
- 通常采用1~2层全连接人工神经网络作分类器



卷积神经网络结构(论文2结构)

▶ 输入层

- 卷积输入层可以直接作用于原始输入数据，对于输入是图像来说，输入数据是图像的像素值

▶ 卷积层

- 卷积神经网络的卷积层，也叫做特征提取层，包括二个部分。
 - 第一部分是真正的卷积层，主要作用是提取输入数据特征。每一个不同的卷积核提取输入数据的特征都不相同，卷积层的卷积核数量越多，就能提取越多输入数据的特征。
 - 第二部分是 pooling 层，也叫下采样层，主要目的是在保留有用信息的基础上减少数据处理量，加快训练网络的速度。通常情况下，卷积神经网络至少包含二层卷积层(这里把真正的卷积层和下采样层统称为卷积层)，即卷积层，pooling层，卷积层，pooling 层。卷积层数越多，在前一层卷积层基础上能够提取更加抽象的特征。

▶ 全连接层

- 可以包含多个全连接层，实际上就是多层感知机的隐含层部分。

▶ 输出层

- 输出层神经节点的数目是根据具体应用任务来设定的。如果是分类任务，卷积神经网络输出层通常是一个分类器

卷积神经网络结构(论文2结构)

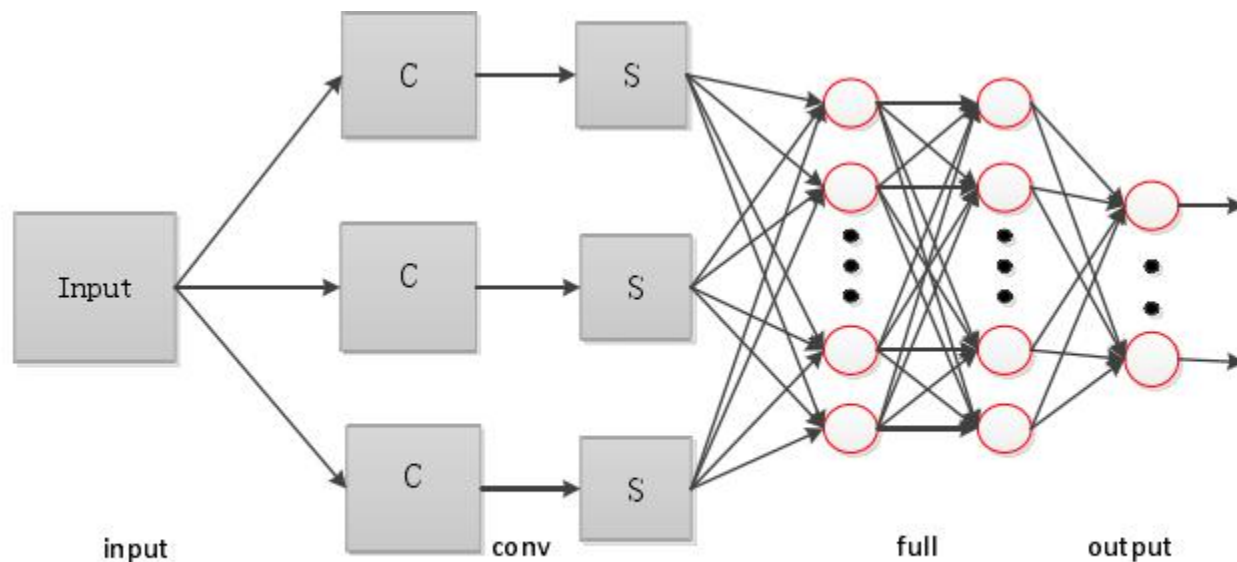


图 2-2 卷积神经网络的基本架构

卷积层的实现

- ▶ 卷积层是卷积神经网络的核心组成部分，其具有局部连接和权值共享特征。
 - 前一层的一个或者多个特征图作为输入与一个或者多个卷积核进行卷积操作，产生一个或者多个输出。
 - 常用的卷积操作为，对于一个大小为 $m*n$ 的特征图，用大小为 $k*k$ 的卷积核对其进行卷积操作，得到的输出特征图的大小为 $(m-k+1)*(n-k+1)$
- ▶ 偏置bias
 - 在卷积操作之后，会在卷积结果上加上一个可训练的参数，称为偏置bias
- ▶ 非线性映射
 - 为了使神经网络具有非线性的拟合性能，须要将得到的结果输入一个非线性的激活函数，通过该函数映射后最终得到卷积层的输出特征图。

卷积层的实现例子

- 5*5的卷积核与输入特征图进行二维离散卷积操作，右边输出特征图中的一个神经元是左侧输入特征图中大小为5*5的连续子区域与卷积核卷积的结果。该子区域就称为该神经元在输入特征图上的感受野，即右边神经元所能“看”到的区域。

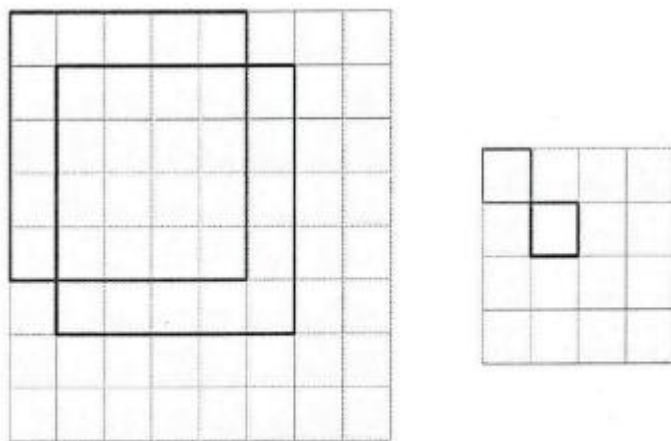
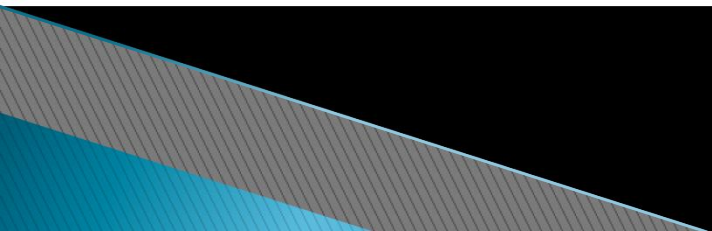


图 2.5 卷积示意图

Fig. 2.5 Diagram of convolution

子采样层的实现

- ▶ 子采样层的作用是对卷积层输出的特征图进行采样，采样层是以采样区域的大小为步长来进行扫描采样，而不是连续的。
- ▶ 采样区域的宽度 w 和高度 h 不一定相等，首先将输入特征图划分为若干个 $w \times h$ 大小的子区域，每个子区域经过子采样之后，对应输出特征图中的一个神经元。



子采样层的实现例子

▶ 子采样层采样模式

- 均值采样：将会计算各个子区域中的均值作为子采样结果；
 - 最大值采样：将会选取各个子区域中的最大值作为子采样结果。
- ▶ 一个大小为 $m*n$ 的输入特征图，经过 $w*h$ 的尺度进行采样之后，得到大小为 $(m/w)*(n/h)$ 的输出特征图。

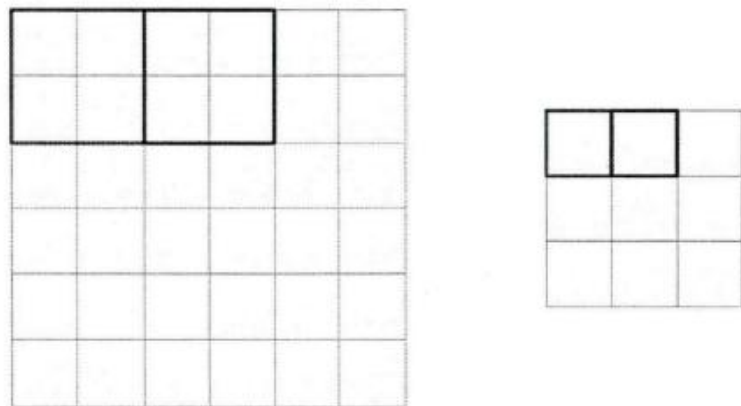


图 2.6 采样示意图

Fig. 2.6 The diagram of sampling

分类器

- ▶ 经过卷积神经网络逐层提取到的特征可以输入任何对于权值可微的分类器。
- ▶ 常用分类器：
 - 1~2层全连接神经网络(默认)
 - 逻辑回归/softmax分类器
 - 支持向量机

卷积神经网络的有监督学习

► 梯度下降法

- 前向传播
 - 依次计算各层的输出值
- 反向传播
 - 依据误差反向依次计算各层权值和偏置的梯度
- 当反向传播结束后，采用梯度下降法进行权值和偏置的权值更新

$$\theta = \theta - \partial \nabla_{\theta} J(\theta; x, y)$$

其中， θ 为需要学习的权值和偏置参数， ∂ 为学习率。

卷积神经网络的无监督学习

- ▶ 思路：用无监督学习获取特征卷积核，利用无监督学习到的特征卷积核取代人工设定的卷积核。
 - 无监督学习运算框架
 - 选择样本：从无标签的训练集图片中，随机提取一些图片
 - 预处理：对这些小的图片采取一些预处理措施
 - 无监督学习映射函数：通过无监督的算法学习到特征函数
 - 当学习到了特征映射函数之后，我们就可以在有标签的数据集上进行特征提取
 - 对于一个输入图像，用上面学习到的特征来与其进行卷积
 - 将上面得到的图像的特征图进行子采样(Pooling)
 - 用第二步得到的特征，结合其原始图片对应的标签一起训练一个线性分类器

The end