

同濟大學

TONGJI UNIVERSITY

毕业设计（论文）

课题名称	人工智能导论报告
副标题	卷积神经网络对应输出数据维度及各层参数量分析
学院	软件学院
专业	软件工程
学生姓名	刘淑仪
学号	2251730
指导教师	曾进
日期	2024 年 4 月 6 日

1 原理分析

1.1 网络层

卷积神经网络由多个不同类型的层组成，每一层都执行特定的数学运算。主要类型包括卷积层（CONV）、激活层（如 Leaky ReLU）、池化层（POOL）、全连接层（FC）等。这些层按顺序堆叠，形成深度学习模型的架构。

1.2 数据维度

在 CNN 中，数据维度通常指的是通过网络流动的多维数组（或张量）的尺寸。例如，一个具有高度（H）、宽度（W）和颜色通道（C）的图像可以表示为一个维度为 [H, W, C] 的张量。当数据通过网络的各层时，它的维度会根据每层的操作而改变。

1.3 参数量

参数是构成深度学习模型的元素，它们在训练过程中被学习和调整。参数量是指模型中总的可学习参数的数量。在卷积层中，参数通常包括卷积核（或过滤器）的权重和偏置项。在全连接层中，参数包括连接权重和偏置项。

1.4 层简介

1.4.1 CONV 层

CONV k-N 表示一个卷积层，其卷积核大小为 $k \times k$ ，输出通道数为 N，此处 padding 为 1，stride（步长）为 1。Padding 是卷积操作中用于保持数据维度的填充，stride 是卷积核移动的步数。

1.4.2 POOL 层

POOL-n 表示一个池化层，这里做 $n \times n$ 大小的池化操作，此处为最大池化，stride 为 n，padding 为 0。池化通常用于降低数据的空间维度。

1.4.3 FC 层

FC-N 表示一个全连接层，其输出节点数为 N。

2 参数分析

2.1 参数分析

1. CONV 3-16：表示一个卷积层，其卷积核大小为 3×3 ，输出通道数为 16。这意味着这层会使用 16 个不同的 3×3 大小的卷积核来处理输入数据。

2. Leaky ReLU : 是一种激活函数, 用于增加网络的非线性。与传统的 ReLU 相比, Leaky ReLU 允许负输入有一个小的正斜率, 从而解决了 ReLU 的“死神经元”问题。

3. POOL-2 : 池化层, 通常用于降低数据的空间维度 (高度和宽度), 从而减少参数量并减轻过拟合。POOL-2 通常指的是使用 2x2 的过滤器进行最大池化。

4. CONV 3-32 : 另一个卷积层, 使用 32 个 3x3 的卷积核来处理输入。

5. FLATTEN : 平坦层, 将前一层的多维输出展平为一维数组, 通常用于卷积层和全连接层之间的过渡。

6. FC-10 : 全连接层, 每个输入节点都与输出层的 10 个节点相连接。这层常用于网络的输出层, 10 可以表示最终的分类数目。

2.2 具体计算方式

卷积核权重的数量: $\text{卷积核权重数量} = \text{输出通道数} \times (\text{输入通道数} \times k \times k)$

偏置项的数量: 输出通道数。

2.3 全连接层参数数量

权重的数量: $\text{权重数量} = (\text{前一层神经元数量}) \times (\text{该层神经元数量})$

偏置项的数量: 该层的神经元数量。

3 结果展示

3.1 分析过程

1. CONV 3-16 : 输出数据维度为 $28 \times 28 \times 16$ 而参数量 (包括 weights 和 biases) 是 448。这里的参数量计算包括了卷积核的权重 ($3 \times 3 \times 3 \times 16$) 和每个卷积核的偏置 (16 个)。

2. Leaky ReLU : 这是一个激活层, 它并不改变数据的维度, 也没有可学习的参数 (即没有权重 weights 和偏置 biases)。故输出数据维度与输入相同, 为 $28 \times 28 \times 16$ 而参数量 (包括 weights 和 biases) 是 0, 因为激活函数不添加额外的参数。

3. POOL-2 : 输出数据维度为 $14 \times 14 \times 16$ 而参数量 (包括 weights 和 biases) 是 0。池化层 (如最大池化层) 不改变数据的深度 (即通道数), 也不包含权重或偏置参数。

4. CONV 3-32 : 输出数据维度为 $14 \times 14 \times 32$ 而参数量 (包括 weights 和 biases) 是 4640。这里的参数量计算包括了卷积核的权重 ($3 \times 3 \times 16 \times 32$) 和每个卷积核的偏置 (32 个)。

5. FLATTEN : 输出数据维度为 6272 (这是将 $14 \times 14 \times 32$ 的三维数据展平成一维向量的结果), 而参数量 (包括 weights 和 biases) 是 0。平坦层 (Flatten) 仅仅是将前面层的多维输出转换为一维数组, 不包含任何参数。

6. FC-10 : 因为全连接层的输出是一个与神经元数量相对应的一维向量, 输出数据维度为 10;

而参数量（包括 weights 和 biases）是 62730。这里的参数量计算包括了连接权重（从 FLATTEN 层的 6272 个输入到 FC 层的 10 个输出，即 6272×10 ）和每个输出神经元的偏置（10 个）。

3.2 结果展示

表 3.1 CNN 层的数据维度和参数量

网络层	数据维度	参数量 (weights+biases)
输入	28×28×3	0
CONV 3-16	28×28×16	448
Leaky ReLU	28×28×16	0
POOL-2	14×14×16	0
CONV 3-32	14×14×32	4640
Leaky ReLU	14×14×32	0
POOL-2	7×7×32	0
FLATTEN	1568	0
FC-10	10	62730

装

订

线