

深度学习

目 录

图 片

表 格

要 点

1 人工智能学家

姓名	图灵奖	诺贝尔奖
Geoffrey Hinton	1	1
John Hopfield	0	1

2 深度学习"深度": 神经网络的深度 (层数)

3 前馈神经网络 (FEED-FORWARD NEURAL NETWORK, FNN)

特点: 相邻层间特征是单向连接。

4 全连接神经网络 (FULL CONNECT NEURAL NETWORK, FCNN)

- 前向传播: 计算结果并保存特征。
- 反向传播: 链式规则。

5 卷积神经网络 (CONVOLUTIONAL NEURAL NETWORKS, CNN)

- 结构
 - 卷积层: 提取局部特征。
 - 池化层: 降低特征维度。
 - 全连接层: 分类。
- 卷积计算过程

原图像大小: $H_0 \times N_0 \times M_0 \times A_0$
卷积核大小: $B \times F \times F \times A_0$

Padding: P

Stride: S

新图像大小: $N_1 = \frac{N_0 + 2P - F}{S} + 1, A_1 = B$

参数量: $F \times F \times A_0 \times B$

- 池化计算过程

最大池化 (Max Pooling): 选择区域的最大值作为代表性的特征值。

平均池化 (Average Pooling): 计算区域的平均值作为代表性的特征值。

池化层参数量为 o 。

- 填充 (Padding)

增加感受野，减少信息损失：确保边缘像素能被卷积核充分覆盖，得到有效处理，而不是丢失。

控制输出尺寸：通过调整填充量可以精确控制每一层的输出尺寸。

- 步幅 (Stride):

控制输出尺寸、下采样程度：较大步幅可以减小输出的空间尺寸，降低计算复杂度，减少参数量，提高特征图的缩放比例。

调整感受野：较大步幅意味着输出单元会覆盖较大输入区域，增加感受野，减少重叠区域数量。

平衡速度与精度：较大步幅可以加速计算过程，但可能丢失细节信息；较小步幅能更精细地捕捉特征，但会增加计算成本。

- 1×1 卷积

维度变换（降维/升维）：改变特征图的深度（通道数），降维有助于降低模型复杂度和计算量，同时保持大部分有用信息。

非线性引入：在卷积后添加激活函数，可以在不改变空间尺寸的情况下引入非线性，使模型能够学习更复杂的模式。

作为瓶颈层：在一些架构中（如 ResNet、Inception），可以用作瓶颈层（先 1×1 卷积降维，再进行其他卷积，最后 1×1 卷积恢复维度）。显著减少参数数量和计算成本，同时维持性能。

特征融合：融合不同尺度或不同来源的特征图，合并成一个新的特征表示。

- 实例

- LeNet: 没有使用ReLU。

- AlexNet: 最早使用了ReLU、GPU。

- VGGNet: 小卷积核 (感受野上, 3个 $3\times 3=1$ 个 7×7)。
- GoogleNet: 使用了ReLU, Inception。 1×1 卷积
- ResNet: 使用了ReLU, 恒等映射直连边, 残差模块。
- 趋势: 卷积核变小、层数增加, 抛弃池化层、全连接层。

6 损失函数

交叉熵损失函数更适用于分类问题, 常用于衡量模型预测的概率分布与真实标记的概率分布之间的差异。