

深度卷积神经网络

一、 Alexnet

计算机视觉研究人员相信,从对最终模型精度的影响来说,更大或更干净的数据集、或是稍微改进的特征提取,比任何学习算法带来的进步要大得多。

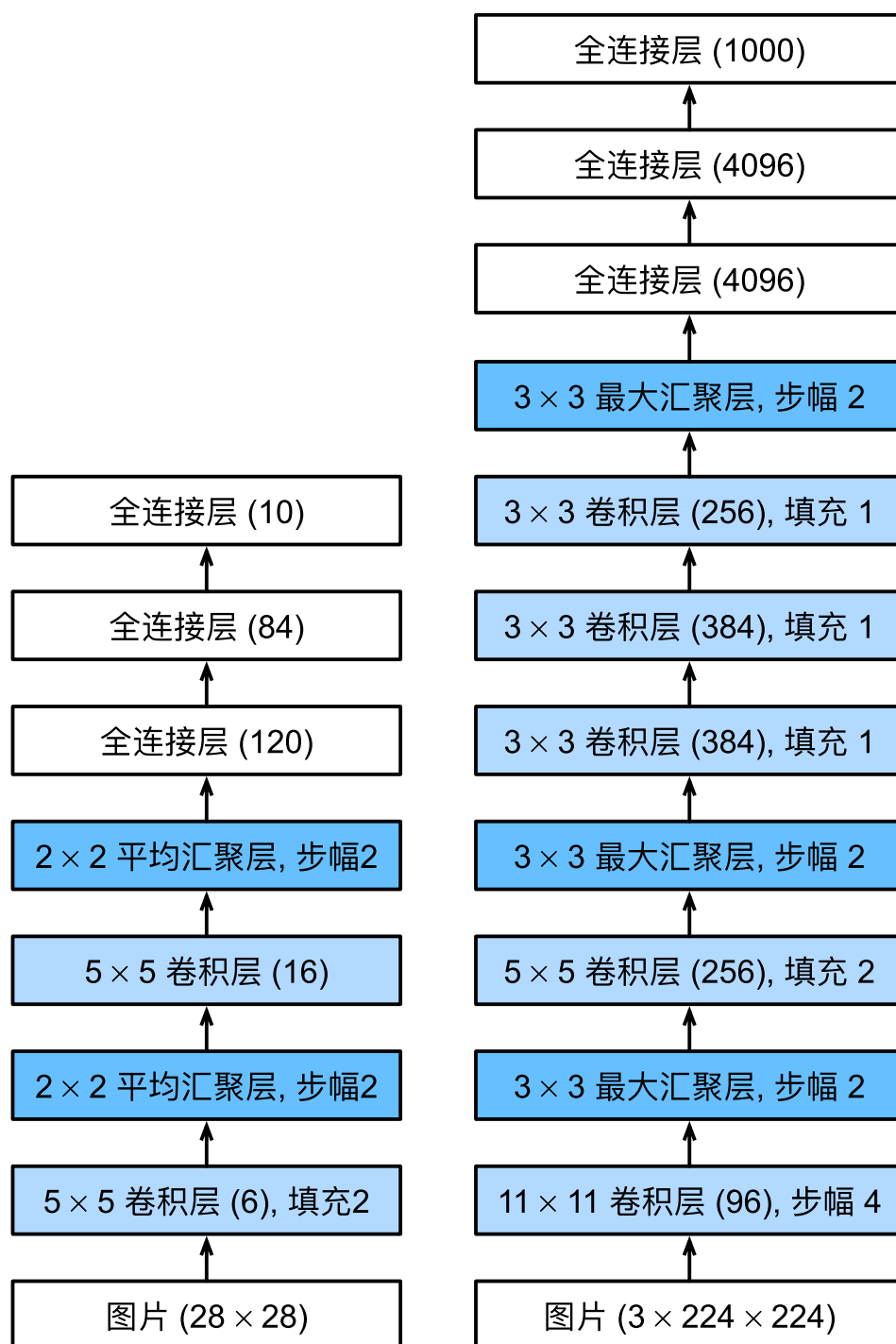


Figure 1 - 1: Alexnet

在 AlexNet 的第一层, 卷积窗口的形状是 11×11 。由于 ImageNet 中大多数图像的宽和高比 MNIST 图像的多 10 倍以上, 因此, 需要一个更大的卷积窗口来捕获

目标。第二层中的卷积窗口形状被缩减为 5×5 ，然后是 3×3 。此外，在第一层、第二层和第五层卷积层之后，加入窗口形状为 3×3 、步幅为 2 的最大汇聚层。而且，AlexNet 的卷积通道数目是 LeNet 的 10 倍。

在最后一个卷积层后有两个全连接层，分别有 4096 个输出。

AlexNet 将 sigmoid 激活函数改为更简单的 ReLU 激活函数。一方面，ReLU 激活函数的计算更简单，它不需要如 sigmoid 激活函数那般复杂的求幂运算。另一方面，当使用不同的参数初始化方法时，ReLU 激活函数使训练模型更加容易。

AlexNet 通过暂退法（4.6 节）控制全连接层的模型复杂度，而 LeNet 只使用了权重衰减。

二、暂退法

一个好的模型需要对输入数据的扰动鲁棒，与之前加入的噪音不一样，之前是固定噪音，丢弃法是随机噪音，丢弃法不是在输入加噪音，而是在层之间加入噪音，所以丢弃法也算是一个正则。

丢弃法对上一层输出向量的每一个元素做如下扰动：

$$x' = \begin{cases} 0 & \text{if 概率为 } p \\ \frac{x}{1-p} & \text{if 其他情况} \end{cases} \quad (2-1)$$

三、引用

[1] 《4.6. 暂退法（Dropout）— 动手学深度学习 2.0.0 documentation》. 见于 2024 年 7 月 7 日. https://zh.d2l.ai/chapter_multilayer-perceptrons/dropout.html#sec-dropout.

[2] 《7.1. 深度卷积神经网络（AlexNet）— 动手学深度学习 2.0.0 documentation》. 见于 2024 年 7 月 7 日. https://zh.d2l.ai/chapter_convolutional-modern/alexnet.html.