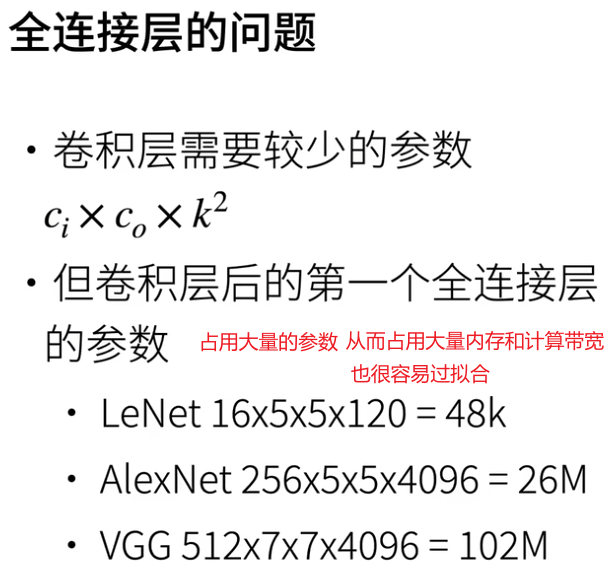
NiN

LeNet、AlexNet和VGG都有一个共同的设计模式：通过一系列的卷积层与汇聚层来提取空间结构特征；然后通过全连接层对特征的表征进行处理。

如果在过程的早期使用全连接层，可能会完全放弃表征的空间结构。

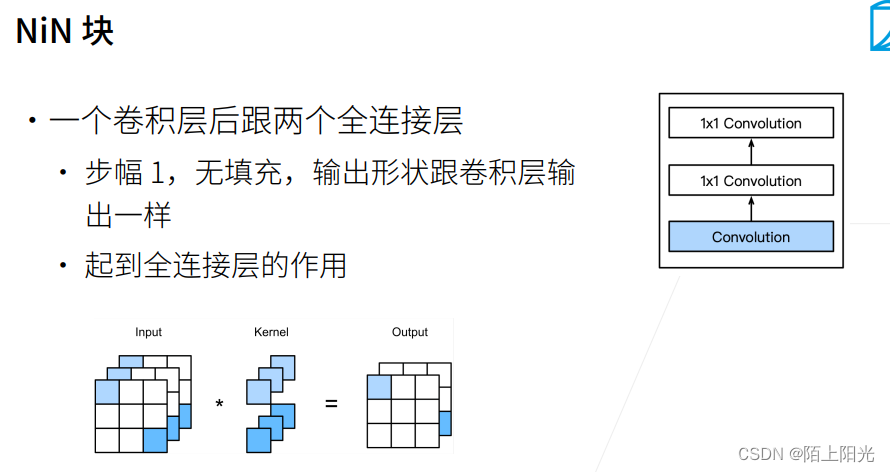
而NiN（网络中的网络）提供了一个非常简单的解决方案：**在每个像素的通道上分别使用多层感知机。**

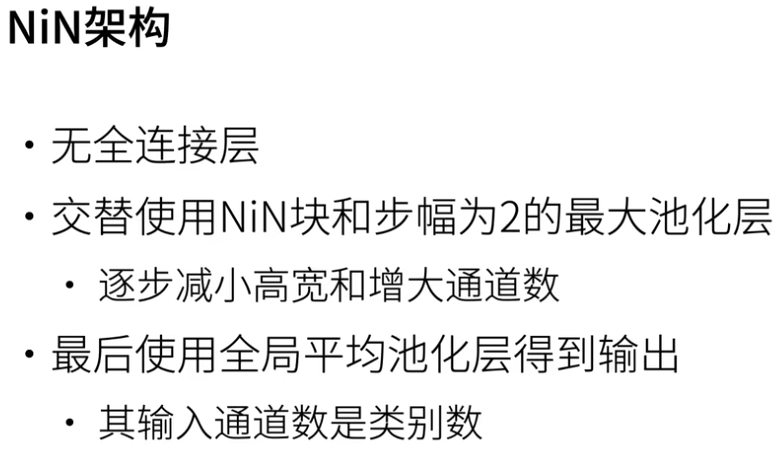


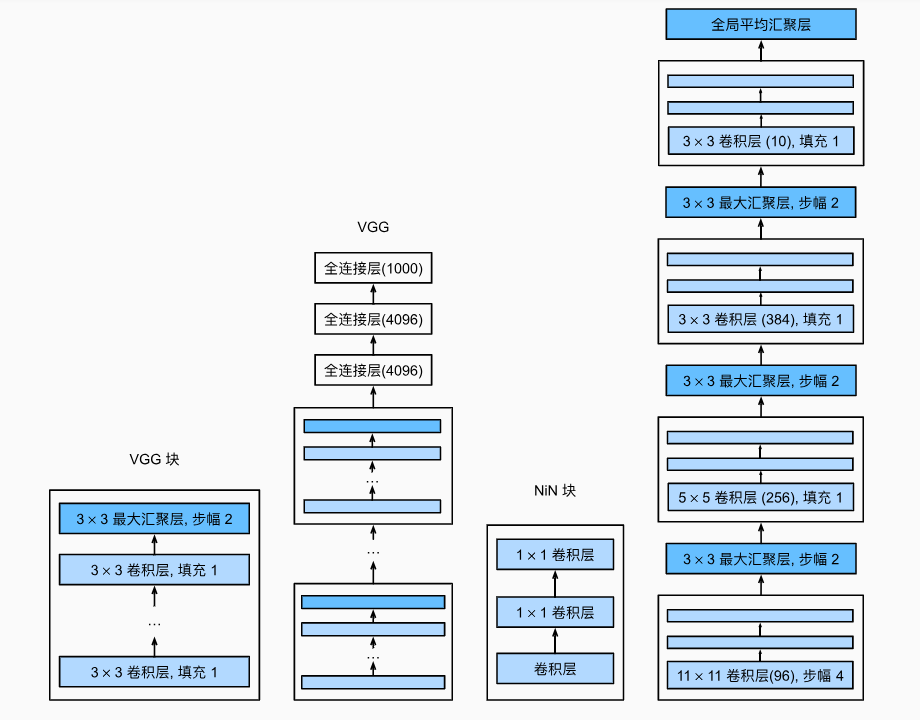
**卷积层**的输入和输出由四维张量组成（样本，通道，高度，宽度）

**全连接层**的输入和输出通常是二维张量（样本，特征）

NiN在每个像素位置（针对每个高度和宽度）应用一个全连接层，可以将其视为1x1卷积层。将间维度中的每个像素视为单个样本，将通道维度视为不同特征。

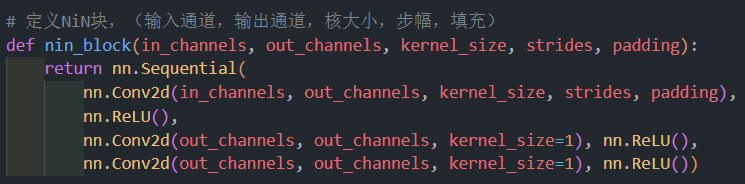






第一层为普通卷积层，之后的两个卷积层充当带有ReLU函数的逐像素全连接层。

代码解读



NiN块由三个卷积层组成，每个卷积层后面都跟着一个ReLU激活函数。第一个卷积层使用较大的卷积核（例如11x11），步幅为4，用于减少输入图像的空间尺寸。接下来的两个卷积层使用1x1的卷积核，步幅为1，用于增加非线性变换的深度。

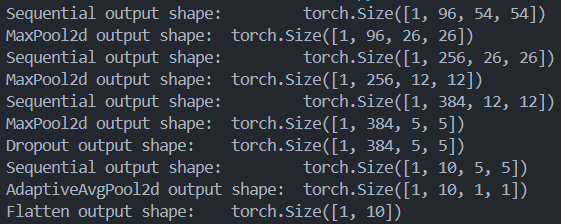


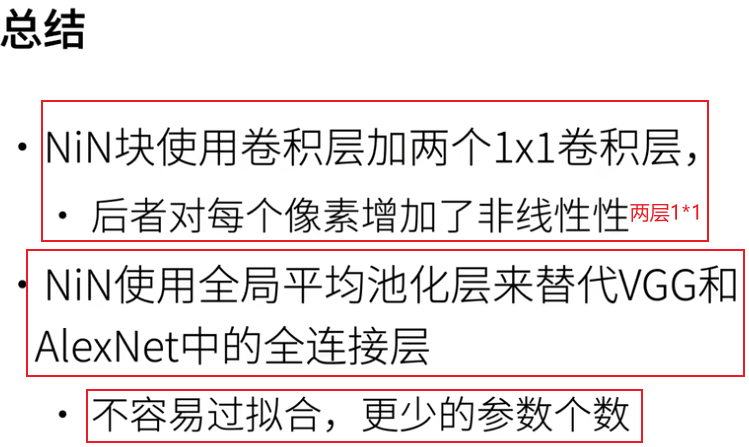
NiN网络架构由多个NiN块和池化层组成（基于AlexNet架构），网络中包含三个NiN块，每个块后面都跟着一个最大池化层（nn.MaxPool2d），用于下采样。第一个NiN块将输入的1个通道（灰度图像）转换为96个通道。在第三个NiN块之后，添加了一个dropout层（nn.Dropout），用于防止过拟合。最后一个NiN块将384个通道转换为10个通道，对应于10个类别（Fashion-MNIST）

**nn.AdaptiveAvgPool2d((1, 1))** 这种操作通常用于将特征图的尺寸统一，以便后续的全连接层（全连接层需要固定大小的输入）—全局平均池化层或者自适应平均池化层

**nn.Flatten()**将四维的输出（批量大小，通道数，1，1）转成二维的输出（批量大小，通道数），其形状为（批量大小，10），然后就可以softmax。

**每个块的输出形状**





Q：卷积层对每个像素增加非线性是什么意思？

A：卷积层，对每个像素对应通道的向量做全连接层，把全连接层对每个像素做连接，2个1\*1卷积层对每个像素的输入通道数做了一个有两个隐藏层的MLP， MLP有ReLU函数，增加了非线性。