

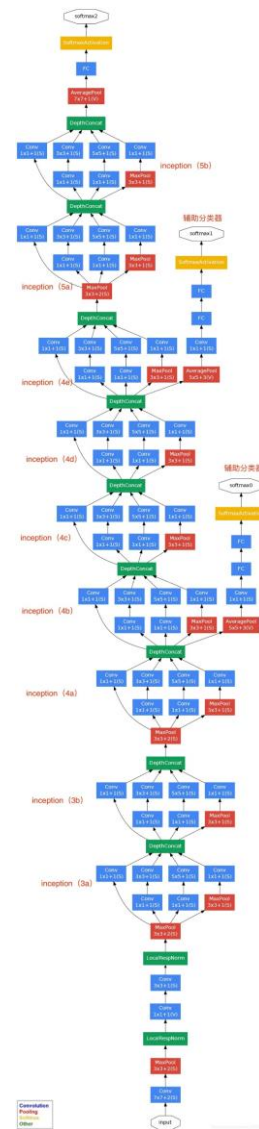
第6章

GoogLeNet

原理与实战

GoogLeNet网络诞生背景

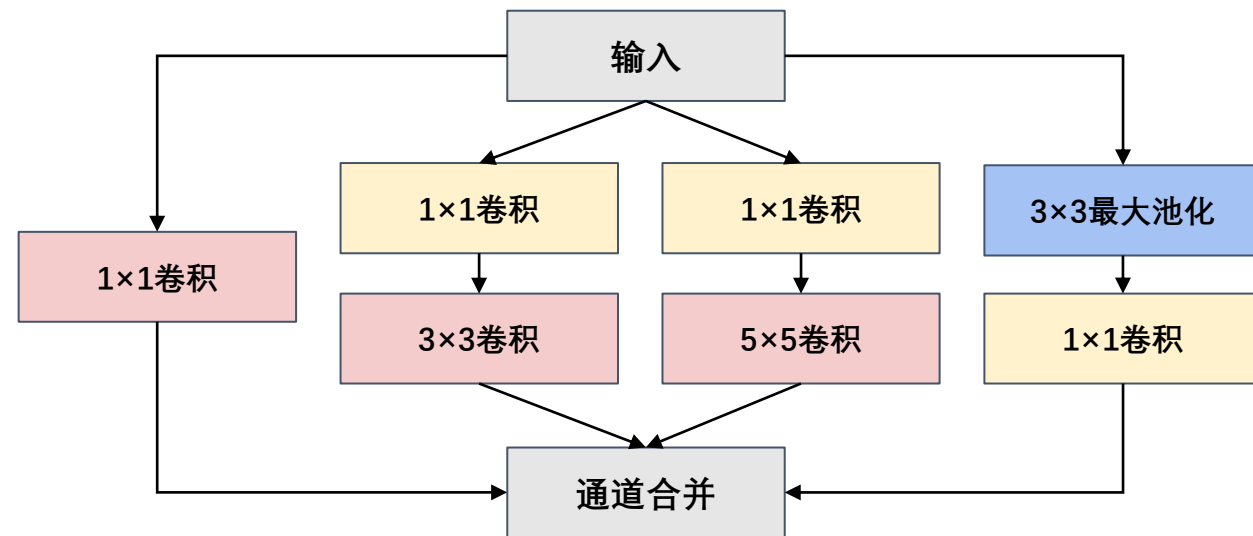
在2014年的ImageNet图像识别挑战赛中，一个名叫 *GoogLeNet* 的网络架构大放异彩。以前流行的网络使用小到 1×1 ，大到 7×7 的卷积核。本文的一个观点是，有时使用不同大小的卷积核组合是有利的。



GoogLeNet网络结构

在GoogLeNet中，基本的卷积块被称为 *Inception块* (Inception block)。这很可能得名于电影《盗梦空间》(Inception)，因为电影中的一句话“我们需要走得更深” (“We need to go deeper”)。

Inception块由四条并行路径组成。前三条路径使用窗口大小为 1×1 、 3×3 和 5×5 的卷积层，从不同空间大小中提取信息。中间的两条路径在输入上执行 1×1 卷积，以减少通道数，从而降低模型的复杂性。第四条路径使用 3×3 最大池化层，然后使用 1×1 卷积层来改变通道数。这四条路径都使用合适的填充来使输入与输出的高和宽一致，最后我们将每条线路的输出在通道维度上连结，并构成Inception块的输出。在Inception块中，通常调整的超参数是每层输出通道数。



GoogLeNet网络结构

输入为 $224 \times 224 \times 3$ 三通道的图像。

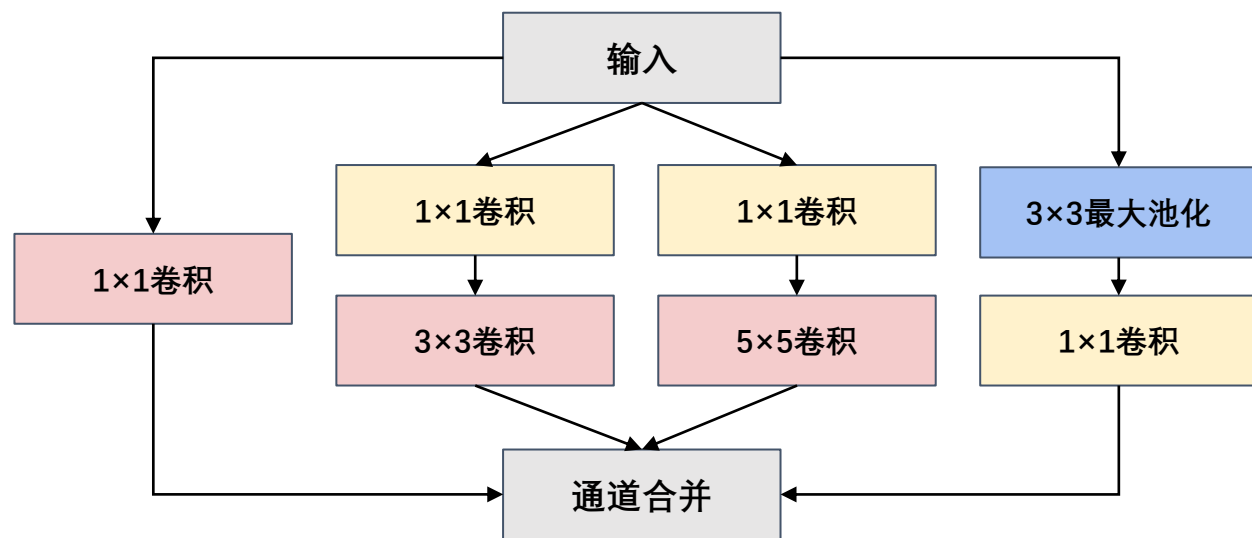
路径1:

(1) 输入为 $224 \times 224 \times 3$ ，卷积核数量为64个；卷积核的尺寸大小为 $1 \times 1 \times 3$ ；步幅为1 (stride=1)，填充为0 (padding=0)；卷积后得到shape为 $224 \times 224 \times 64$ 的特征图输出。

路径2:

(1) 输入为 $224 \times 224 \times 3$ ，卷积核数量为96个；卷积核的尺寸大小为 $1 \times 1 \times 3$ ；步幅为1 (stride = 1)，填充为0 (padding=0)；卷积后得到shape为 $224 \times 224 \times 64$ 的特征图输出。

(2) 输入为 $224 \times 224 \times 64$ ，卷积核数量为128个；卷积核的尺寸大小为 $3 \times 3 \times 64$ ；步幅为1 (stride = 1)，填充为1 (padding=1)；卷积后得到shape为 $224 \times 224 \times 128$ 的特征图输出。



GoogLeNet网络结构

路径3:

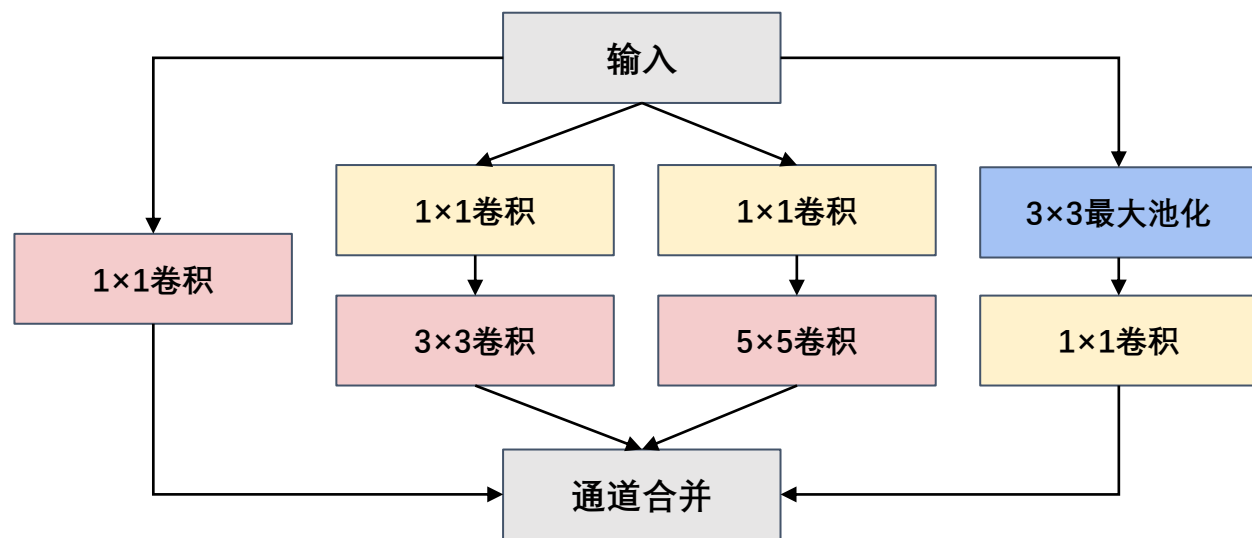
(1) 输入为 $224 \times 224 \times 3$ ，卷积核数量为16个；卷积核的尺寸大小为 $1 \times 1 \times 3$ ；步幅为1 (stride = 1)，填充为0 (padding=0)；卷积后得到shape为 $224 \times 224 \times 16$ 的特征图输出。

(2) 输入为 $224 \times 224 \times 16$ ，卷积核数量为32个；卷积核的尺寸大小为 $5 \times 5 \times 16$ ；步幅为1 (stride = 1)，填充为2 (padding=2)；卷积后得到shape为 $224 \times 224 \times 32$ 的特征图输出。

路径4:

(1) 输入为 $224 \times 224 \times 3$ ，池化核的尺寸大小为 3×3 ；步幅为1 (stride = 1)，填充为1 (padding=1)；池化后得到shape为 $224 \times 224 \times 3$ 的特征图输出。

(2) 输入为 $224 \times 224 \times 3$ ，卷积核数量为32个；卷积核的尺寸大小为 $1 \times 1 \times 3$ ；步幅为1 (stride = 1)，填充为0 (padding=0)；卷积后得到shape为 $224 \times 224 \times 32$ 的特征图输出。



GoogLeNet网络结构

通道合并:

路径1的到输出为: $224 \times 224 \times 64$

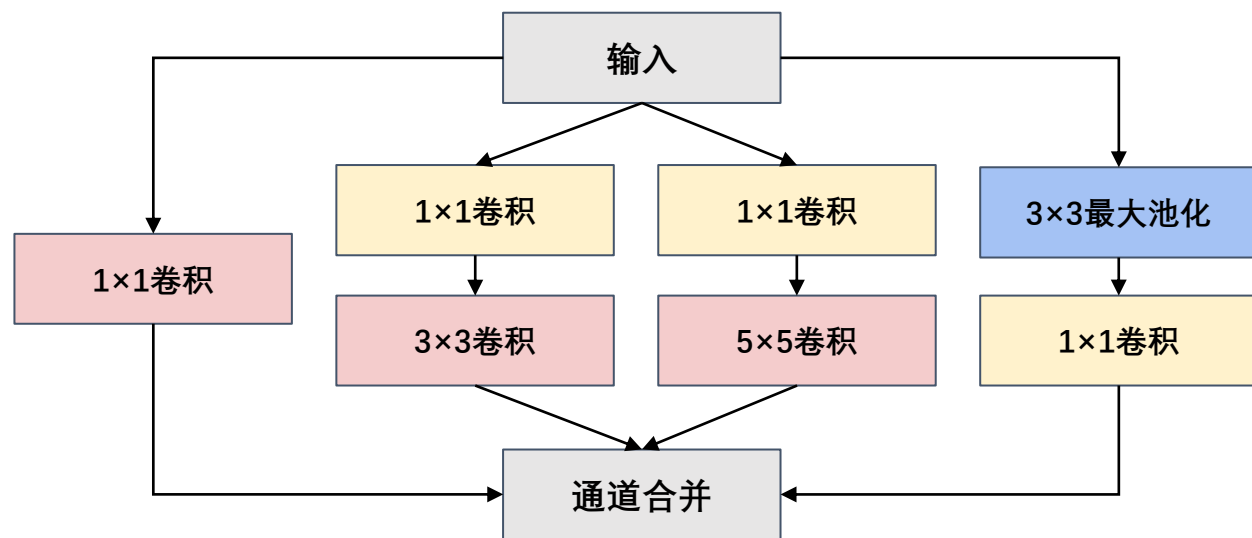
路径2的到输出为: $224 \times 224 \times 128$

路径3的到输出为: $224 \times 224 \times 32$

路径4的到输出为: $224 \times 224 \times 32$

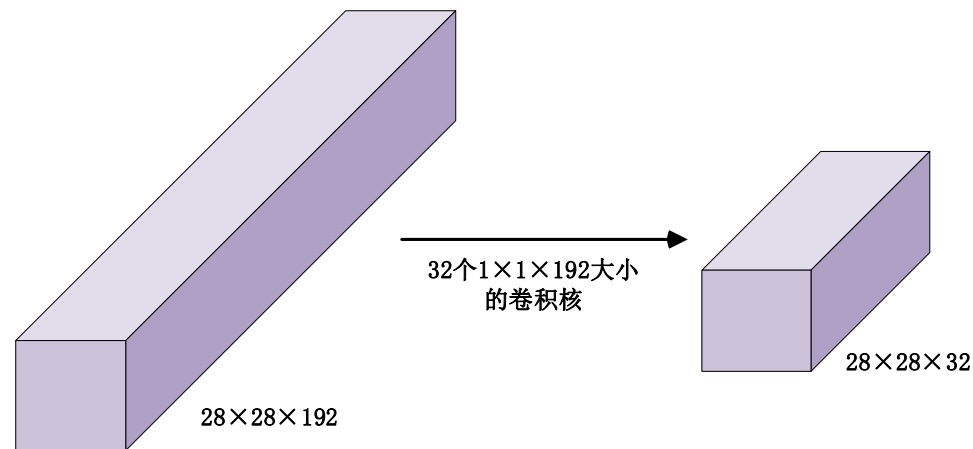
最终通道合并为 $64 + 128 + 32 + 32 = 256$, 最终的输出为: $224 \times 224 \times 256$ 。

那么为什么GoogLeNet这个网络如此有效呢? 首先我们考虑一下滤波器 (filter) 的组合, 它们可以用各种滤波器尺寸探索图像, 这意味着不同大小的滤波器可以有效地识别不同范围的图像细节。同时, 我们可以为不同的滤波器分配不同数量的参数。

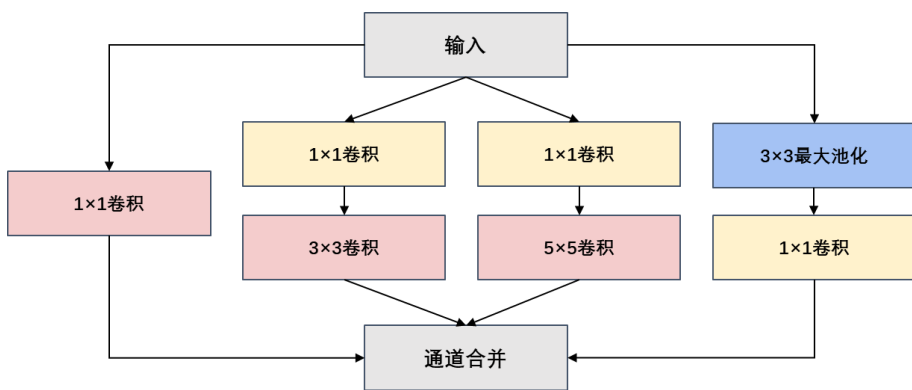


1*1卷积的优点

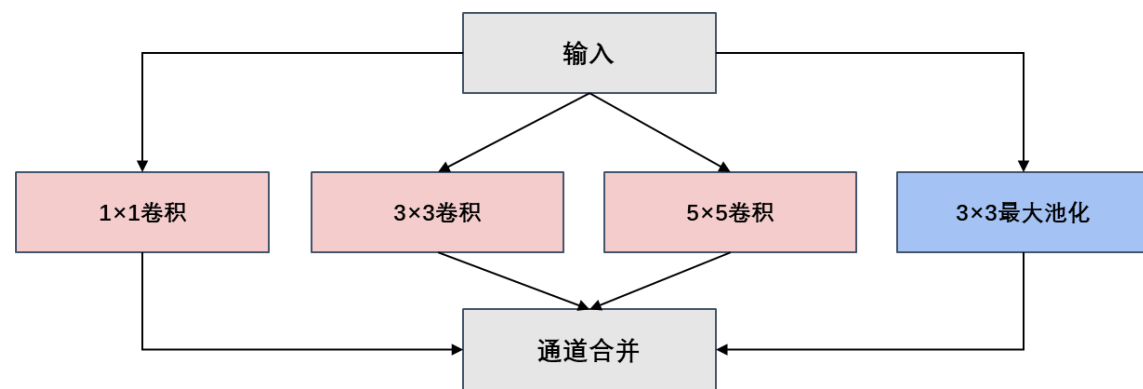
1*1卷积的作用主要是：
实现跨通道的交互和信息整合
卷积核通道数的降维和升维，减少网络参数



1*1卷积的优点



如图中输入的特征图大小为 $(28 \times 28 \times 192)$ ，经过 1×1 卷积后输出的特征图大小为 $28 \times 28 \times 96$ ，再将 1×1 卷积后的特征图送入 3×3 卷积中，得到特征图的通道大小为 128，因此改路径的参数量为： $(1 \times 1 \times 192 \times 96) + (3 \times 3 \times 96 \times 128) = 129024$ 。对比可知，加入 1×1 卷积后参数量减少了。

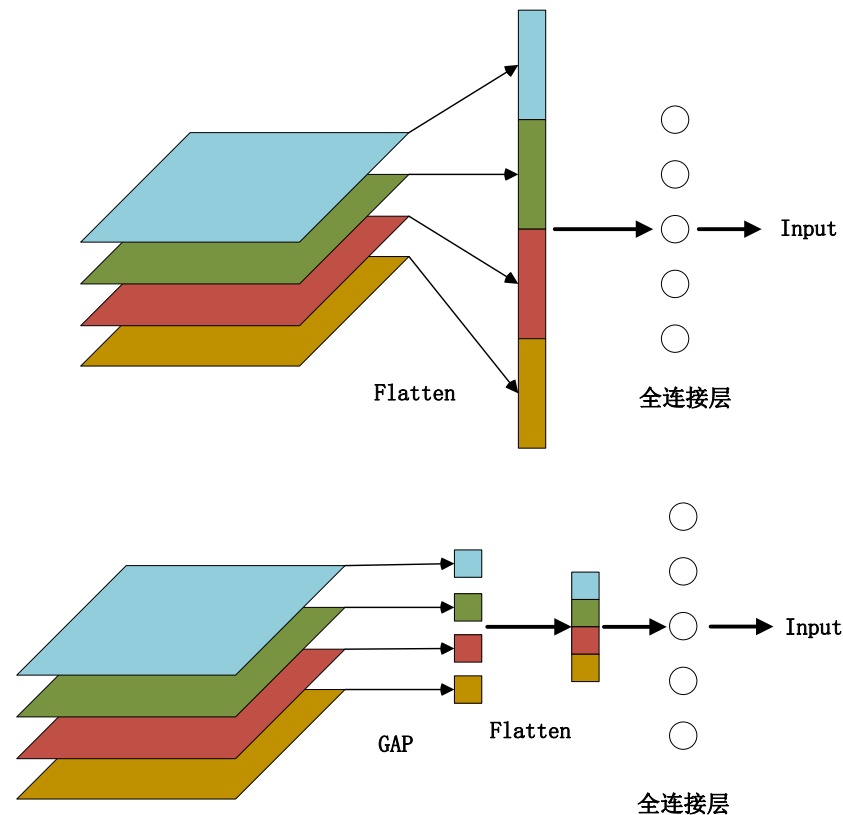


我们以 3×3 卷积线路为例，假设输入的特征图大小为 $(28 \times 28 \times 192)$ ，输出特征图的通道数是 128：如图中该线路的参数量为： $3 \times 3 \times 192 \times 128 = 221184$

全局平均池化层（GAP） --优势

抑制过拟合：直接拉平做全连接层的方式依然保留了大量的空间信息，假设feature map是32个通道的10*10图像，那么拉平就得到了32*10*10的向量，如果是最后一层是对应两类标签，那么这一层就需要3200*2的权重矩阵，而GAP不同，将空间上的信息直接用均值代替，32个通道GAP之后得到的向量都是32的向量，那么最后一层只需要32*2的权重矩阵。相比之下GAP网络参数会更少，而全连接更容易在大量保留下来的空间信息上面过拟合。

GAP使特征图输入尺寸更加灵活：在前面举例里面可以看到特征图经过GAP后的大小为 $1 \times 1 \times C$ ，这里的C为通道的数量，因此，此刻神经网络参数不再与输入图像尺寸的大小有关，和输入特征图的通道有关，也就是输入图像的长宽可以不固定。



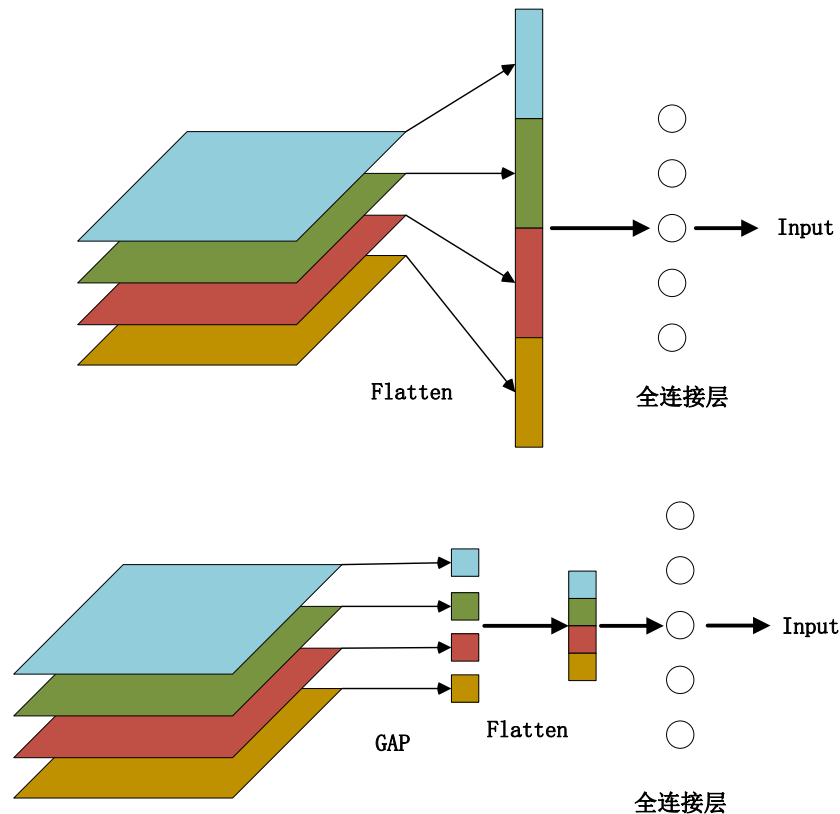
全局平均池化层（GAP）--缺点

信息丢失：全局平均池化会将每个通道内的特征信息压缩成一个单一的数值。这可能会导致一些特征信息的丢失，特别是在深层网络中，重要的细节可能会被平均掉，从而降低了网络的表达能力。

特征丰富性：在一些任务中，特征之间的关系和细节对于正确的分类或预测是至关重要的。全局平均池化可能无法很好地捕捉这些特征之间的复杂关系，从而影响了模型的性能。

梯度信息：在深度学习中，梯度信息是用于权重更新的关键。在全局平均池化之后，特征图被降维为一个向量，导致梯度的传播变得更加困难。这可能会导致梯度消失或爆炸，从而影响了优化过程。

复杂任务：对于一些复杂的任务，模型需要更多的层级和特征表示，而全局平均池化可能无法提供足够的表达能力。



GoogLeNet网络参数详解

第1层输入层： 输入为 $224 \times 224 \times 3$ 三通道的图像。

第2块：

(1) 卷积层；输入为 $224 \times 224 \times 3$ ，卷积核数量为64个；卷积核的尺寸大小为 $7 \times 7 \times 3$ ；步幅为2（ $\text{stride} = 2$ ），填充为3（ $\text{padding}=3$ ）；卷积后得到shape为 $112 \times 112 \times 64$ 的特征图输出。

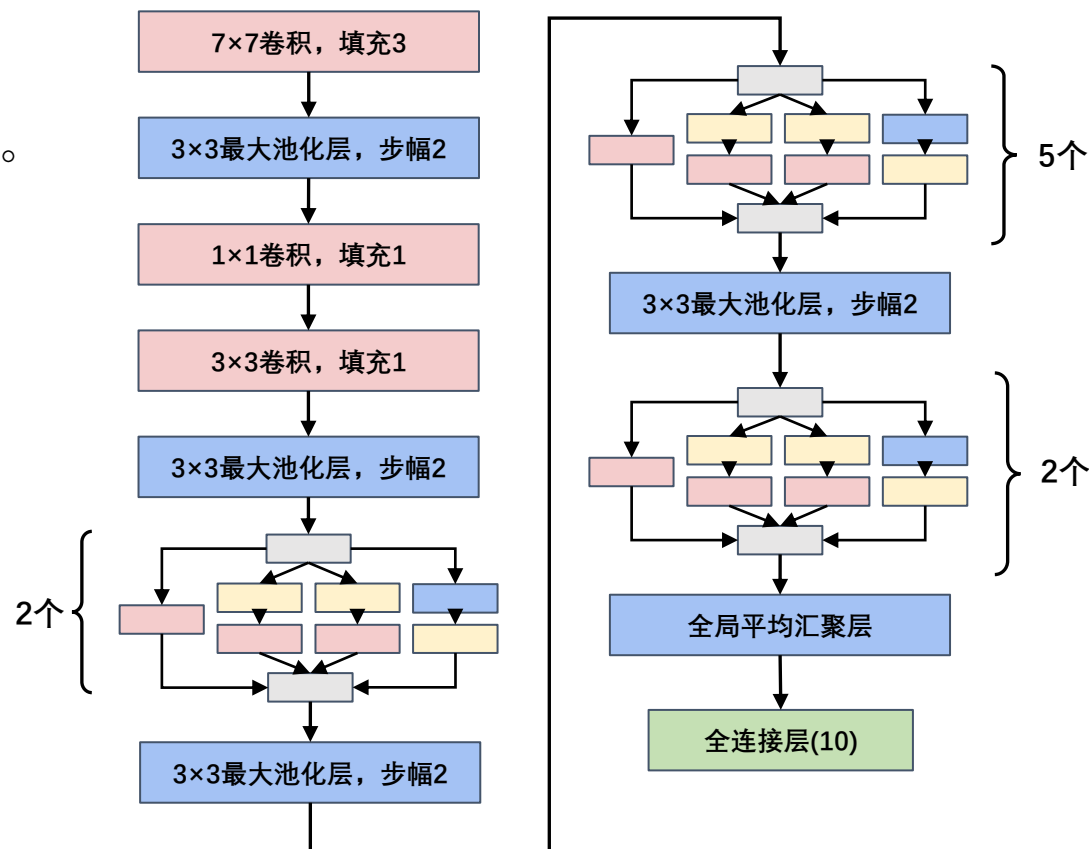
(2) 最大池化层；输入为 $112 \times 112 \times 64$ ，池化核为 3×3 ，步幅为2（ $\text{stride} = 2$ ），填充为1（ $\text{padding}=1$ ）；后得到尺寸为 $56 \times 56 \times 64$ 的池化层的特征图输出。

第3块：

(1) 卷积层；输入为 $56 \times 56 \times 64$ ，卷积核数量为64个；卷积核的尺寸大小为 $1 \times 1 \times 64$ ；步幅为1（ $\text{stride} = 1$ ），填充为0（ $\text{padding}=0$ ）；卷积后得到shape为 $56 \times 56 \times 64$ 的特征图输出。

(2) 卷积层；输入为 $56 \times 56 \times 64$ ，卷积核数量为192个；卷积核的尺寸大小为 $3 \times 3 \times 64$ ；步幅为1（ $\text{stride} = 1$ ），填充为1（ $\text{padding}=1$ ）；卷积后得到shape为 $56 \times 56 \times 192$ 的特征图输出。

(3) 最大池化层；输入为 $56 \times 56 \times 192$ ，池化核为 3×3 ，步幅为2（ $\text{stride} = 2$ ），填充为1（ $\text{padding}=1$ ）；后得到尺寸为 $28 \times 28 \times 192$ 的池化层的特征图输出。



GoogLeNet网络参数详解--第1个Inception块

第1个Inception块:

输入为 $28 \times 28 \times 192$ 的特征图。

路径1:

(1) 输入为 $28 \times 28 \times 192$ ，卷积核数量为64个；卷积核的尺寸大小为 $1 \times 1 \times 192$ ；步幅为1 (stride = 1)，填充为0 (padding=0)；卷积后得到shape为 $28 \times 28 \times 64$ 的特征图输出。

路径2:

(1) 输入为 $28 \times 28 \times 192$ ，卷积核数量为96个；卷积核的尺寸大小为 $1 \times 1 \times 192$ ；步幅为1 (stride = 1)，填充为0 (padding=0)；卷积后得到shape为 $28 \times 28 \times 96$ 的特征图输出。

(2) 输入为 $28 \times 28 \times 96$ ，卷积核数量为128个；卷积核的尺寸大小为 $3 \times 3 \times 96$ ；步幅为1 (stride = 1)，填充为1 (padding=1)；卷积后得到shape为 $28 \times 28 \times 128$ 的特征图输出。

路径3:

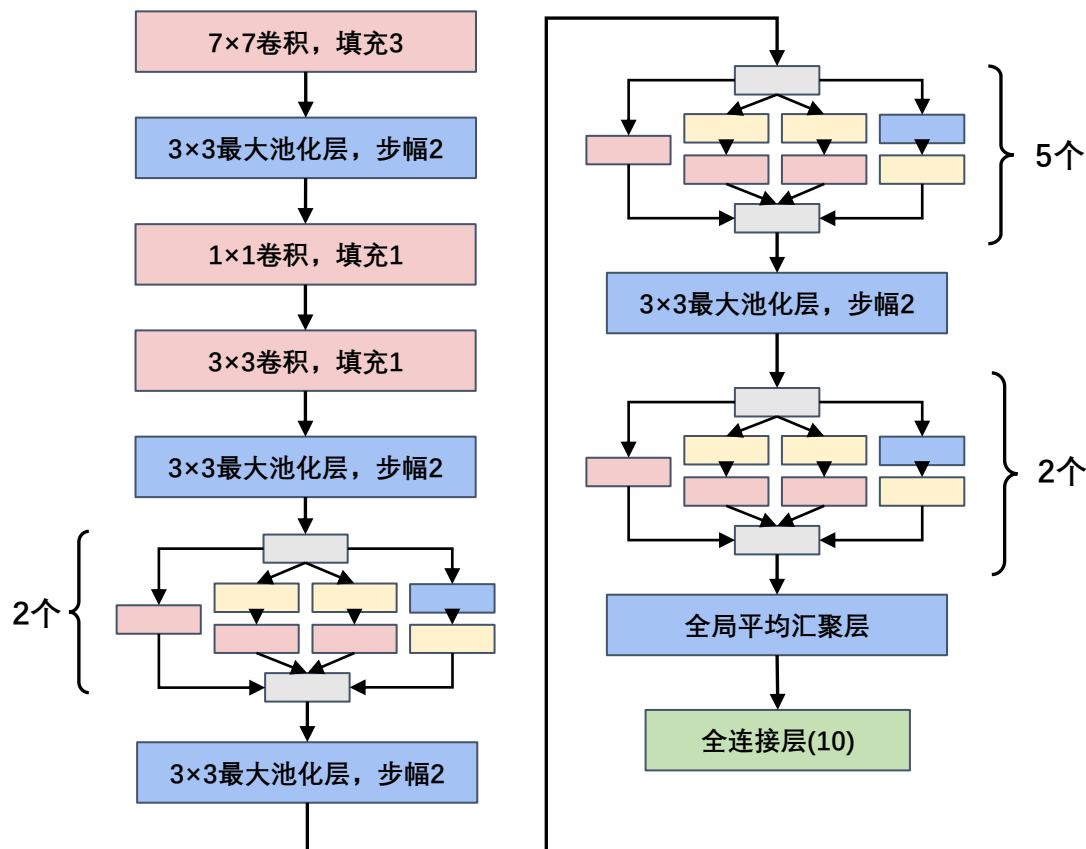
(1) 输入为 $28 \times 28 \times 192$ ，卷积核数量为16个；卷积核的尺寸大小为 $1 \times 1 \times 192$ ；步幅为1 (stride = 1)，填充为0 (padding=0)；卷积后得到shape为 $28 \times 28 \times 16$ 的特征图输出。

(2) 输入为 $28 \times 28 \times 16$ ，卷积核数量为32个；卷积核的尺寸大小为 $5 \times 5 \times 16$ ；步幅为1 (stride = 1)，填充为2 (padding=2)；卷积后得到shape为 $28 \times 28 \times 32$ 的特征图输出。

路径4:

(1) 输入为 $28 \times 28 \times 192$ ，池化核的尺寸大小为 3×3 ；步幅为1 (stride = 1)，填充为1 (padding=1)；池化后得到shape为 $28 \times 28 \times 192$ 的特征图输出。

(2) 输入为 $28 \times 28 \times 192$ ，卷积核数量为32个；卷积核的尺寸大小为 $1 \times 1 \times 3$ ；步幅为1 (stride = 1)，填充为0 (padding=0)；卷积后得到shape为 $28 \times 28 \times 32$ 的特征图输出。



GoogLeNet网络参数详解--第1个Inception块

通道合并:

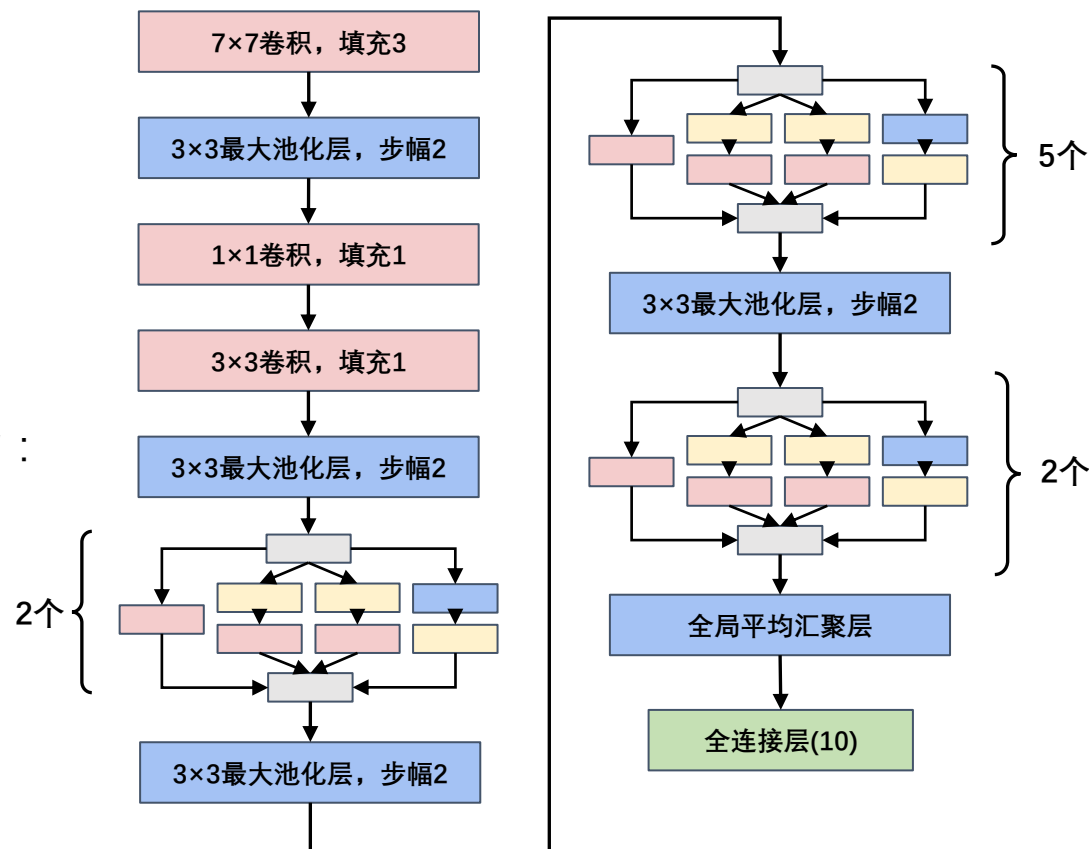
路径1的到输出为: $28 \times 28 \times 64$

路径2的到输出为: $28 \times 28 \times 128$

路径3的到输出为: $28 \times 28 \times 32$

路径4的到输出为: $28 \times 28 \times 32$

最终通道合并为 $64+128+32+32=256$, 最终的输出为:
 $28 \times 28 \times 256$ 。



GoogLeNet网络参数详解--第2个Inception块

第2个Inception块:

输入为 $28 \times 28 \times 256$ 的特征图。

路径1:

(1) 输入为 $28 \times 28 \times 256$ ，卷积核数量为128个；卷积核的尺寸大小为 $1 \times 1 \times 256$ ；步幅为1 (stride = 1)，填充为0 (padding=0)；卷积后得到shape为 $28 \times 28 \times 128$ 的特征图输出。

路径2:

(1) 输入为 $28 \times 28 \times 256$ ，卷积核数量为128个；卷积核的尺寸大小为 $1 \times 1 \times 256$ ；步幅为1 (stride = 1)，填充为0 (padding=0)；卷积后得到shape为 $28 \times 28 \times 128$ 的特征图输出。

(2) 输入为 $28 \times 28 \times 256$ ，卷积核数量为192个；卷积核的尺寸大小为 $3 \times 3 \times 256$ ；步幅为1 (stride = 1)，填充为1 (padding=1)；卷积后得到shape为 $28 \times 28 \times 192$ 的特征图输出。

路径3:

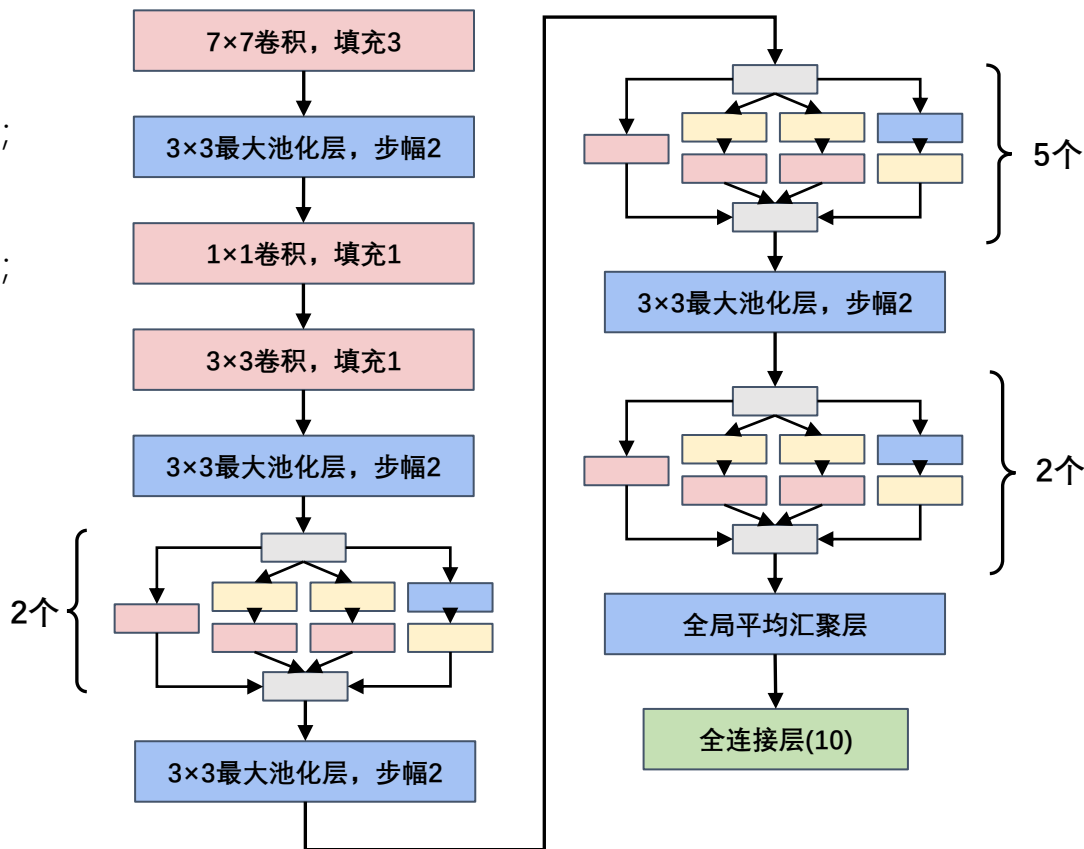
(1) 输入为 $28 \times 28 \times 256$ ，卷积核数量为32个；卷积核的尺寸大小为 $1 \times 1 \times 256$ ；步幅为1 (stride = 1)，填充为0 (padding=0)；卷积后得到shape为 $28 \times 28 \times 32$ 的特征图输出。

(2) 输入为 $28 \times 28 \times 32$ ，卷积核数量为96个；卷积核的尺寸大小为 $5 \times 5 \times 32$ ；步幅为1 (stride = 1)，填充为2 (padding=2)；卷积后得到shape为 $28 \times 28 \times 96$ 的特征图输出。

路径4:

(1) 输入为 $28 \times 28 \times 256$ ，池化核的尺寸大小为 3×3 ；步幅为1 (stride = 1)，填充为1 (padding=1)；池化后得到shape为 $28 \times 28 \times 256$ 的特征图输出。

(2) 输入为 $28 \times 28 \times 256$ ，卷积核数量为64个；卷积核的尺寸大小为 $1 \times 1 \times 256$ ；步幅为1 (stride = 1)，填充为0 (padding=0)；卷积后得到shape为 $28 \times 28 \times 64$ 的特征图输出。



GoogLeNet网络参数详解--第2个Inception块

通道合并:

路径1的到输出为: $28 \times 28 \times 128$

路径2的到输出为: $28 \times 28 \times 192$

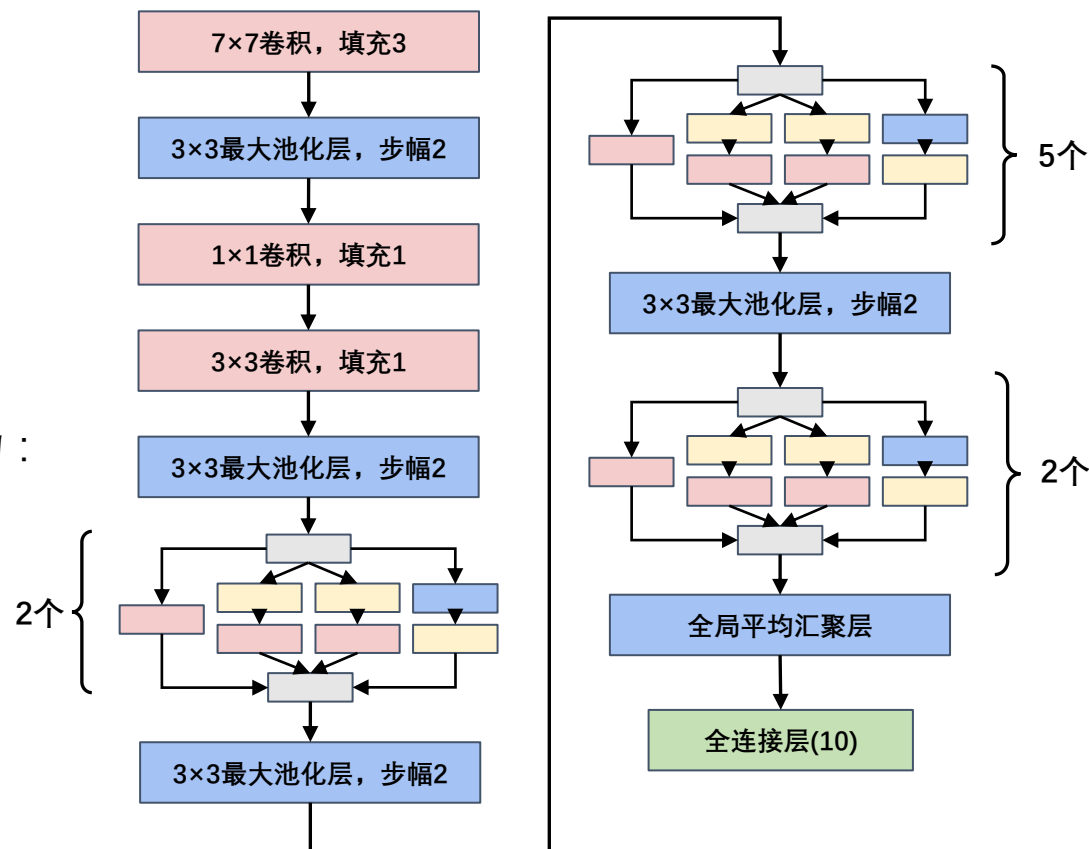
路径3的到输出为: $28 \times 28 \times 96$

路径4的到输出为: $28 \times 28 \times 64$

最终通道合并为 $128+192+96+64=480$, 最终的输出为:
 $28 \times 28 \times 480$ 。

最大池化模块:

输入为 $28 \times 28 \times 480$ 。池化核的尺寸大小为 3×3 ; 步幅为2 (stride = 2), 填充为1 (padding=1); 池化后得到shape为 $14 \times 14 \times 480$ 的特征图输出。



GoogLeNet网络参数详解--第3个Inception块

第3个Inception块:

输入为 $14 \times 14 \times 480$ 的特征图。

路径1:

(1) 输入为 $14 \times 14 \times 480$ ，卷积核数量为192个；卷积核的尺寸大小为 $1 \times 1 \times 480$ ；步幅为1 (stride = 1)，填充为0 (padding=0)；卷积后得到shape为 $14 \times 14 \times 192$ 的特征图输出。

路径2:

(1) 输入为 $14 \times 14 \times 480$ ，卷积核数量为96个；卷积核的尺寸大小为 $1 \times 1 \times 480$ ；步幅为1 (stride = 1)，填充为0 (padding=0)；卷积后得到shape为 $14 \times 14 \times 96$ 的特征图输出。

(2) 输入为 $14 \times 14 \times 96$ ，卷积核数量为208个；卷积核的尺寸大小为 $3 \times 3 \times 96$ ；步幅为1 (stride = 1)，填充为1 (padding=1)；卷积后得到shape为 $14 \times 14 \times 208$ 的特征图输出。

路径3:

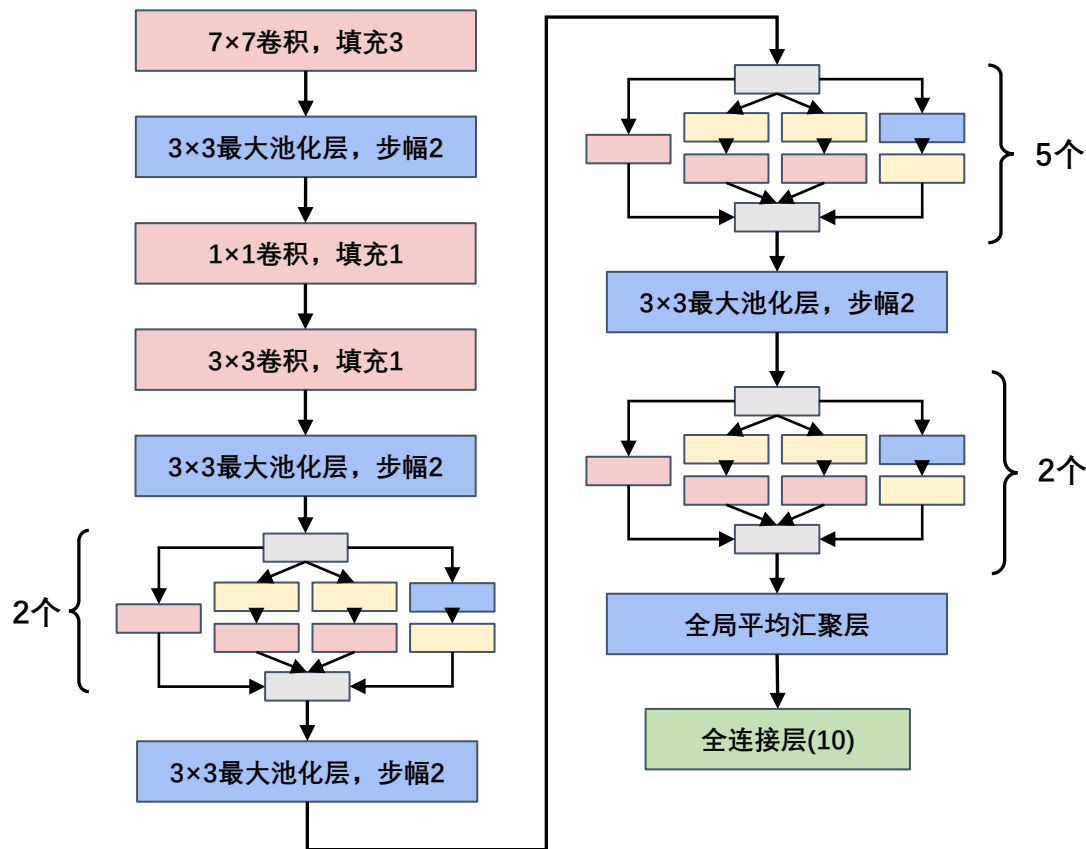
(1) 输入为 $14 \times 14 \times 480$ ，卷积核数量为16个；卷积核的尺寸大小为 $1 \times 1 \times 480$ ；步幅为1 (stride = 1)，填充为0 (padding=0)；卷积后得到shape为 $14 \times 14 \times 16$ 的特征图输出。

(2) 输入为 $14 \times 14 \times 16$ ，卷积核数量为48个；卷积核的尺寸大小为 $5 \times 5 \times 16$ ；步幅为1 (stride = 1)，填充为2 (padding=2)；卷积后得到shape为 $14 \times 14 \times 48$ 的特征图输出。

路径4:

(1) 输入为 $14 \times 14 \times 480$ ，池化核的尺寸大小为 3×3 ；步幅为1 (stride = 1)，填充为1 (padding=1)；池化后得到shape为 $14 \times 14 \times 480$ 的特征图输出。

(2) 输入为 $14 \times 14 \times 480$ ，卷积核数量为64个；卷积核的尺寸大小为 $1 \times 1 \times 3$ ；步幅为1 (stride = 1)，填充为0 (padding=0)；卷积后得到shape为 $14 \times 14 \times 64$ 的特征图输出。



GoogLeNet网络参数详解--第3个Inception块

通道合并:

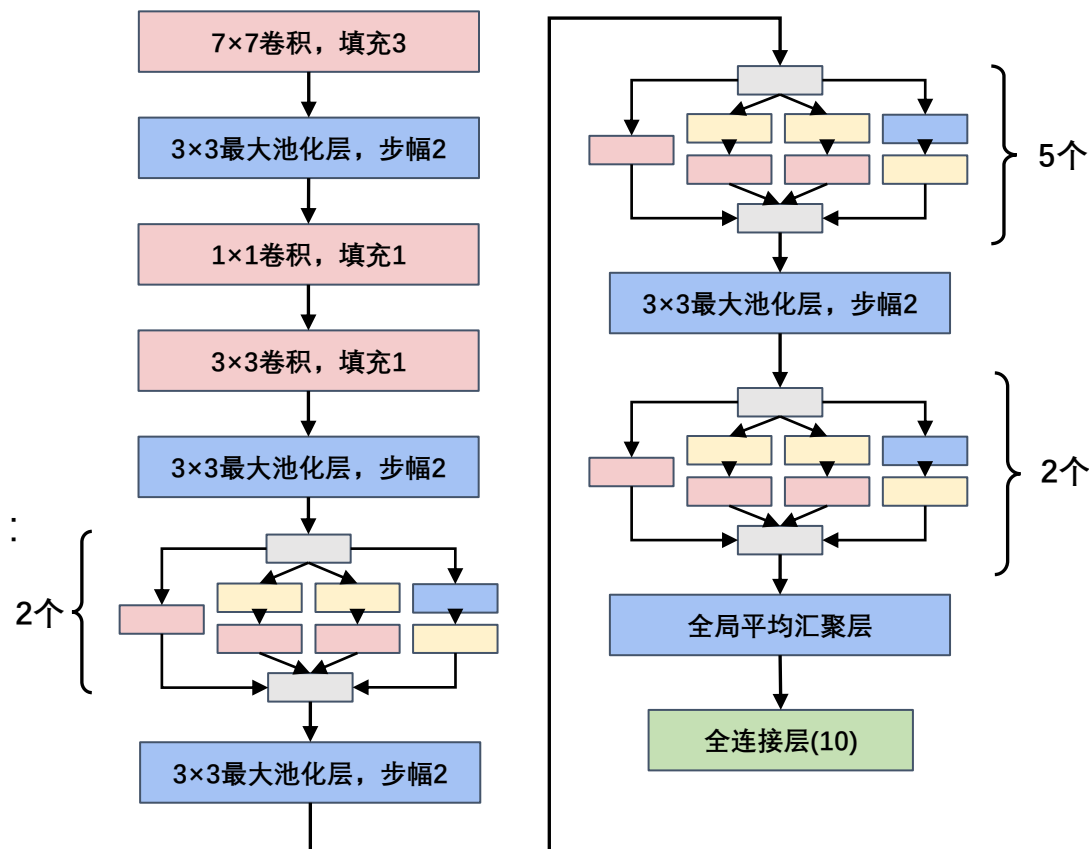
路径1的到输出为: $14 \times 14 \times 192$

路径2的到输出为: $14 \times 14 \times 208$

路径3的到输出为: $14 \times 14 \times 48$

路径4的到输出为: $14 \times 14 \times 64$

最终通道合并为 $192 + 208 + 48 + 64 = 512$, 最终的输出为:
 $14 \times 14 \times 512$ 。



GoogLeNet网络参数详解--第4个Inception块

第3个Inception块:

输入为 $14 \times 14 \times 512$ 的特征图。

路径1:

(1) 输入为 $14 \times 14 \times 512$ ，卷积核数量为160个；卷积核的尺寸大小为 $1 \times 1 \times 512$ ；步幅为1 (stride = 1)，填充为0 (padding=0)；卷积后得到shape为 $14 \times 14 \times 160$ 的特征图输出。

路径2:

(1) 输入为 $14 \times 14 \times 512$ ，卷积核数量为112个；卷积核的尺寸大小为 $1 \times 1 \times 512$ ；步幅为1 (stride = 1)，填充为0 (padding=0)；卷积后得到shape为 $14 \times 14 \times 112$ 的特征图输出。

(2) 输入为 $14 \times 14 \times 112$ ，卷积核数量为224个；卷积核的尺寸大小为 $3 \times 3 \times 112$ ；步幅为1 (stride = 1)，填充为1 (padding=1)；卷积后得到shape为 $14 \times 14 \times 224$ 的特征图输出。

路径3:

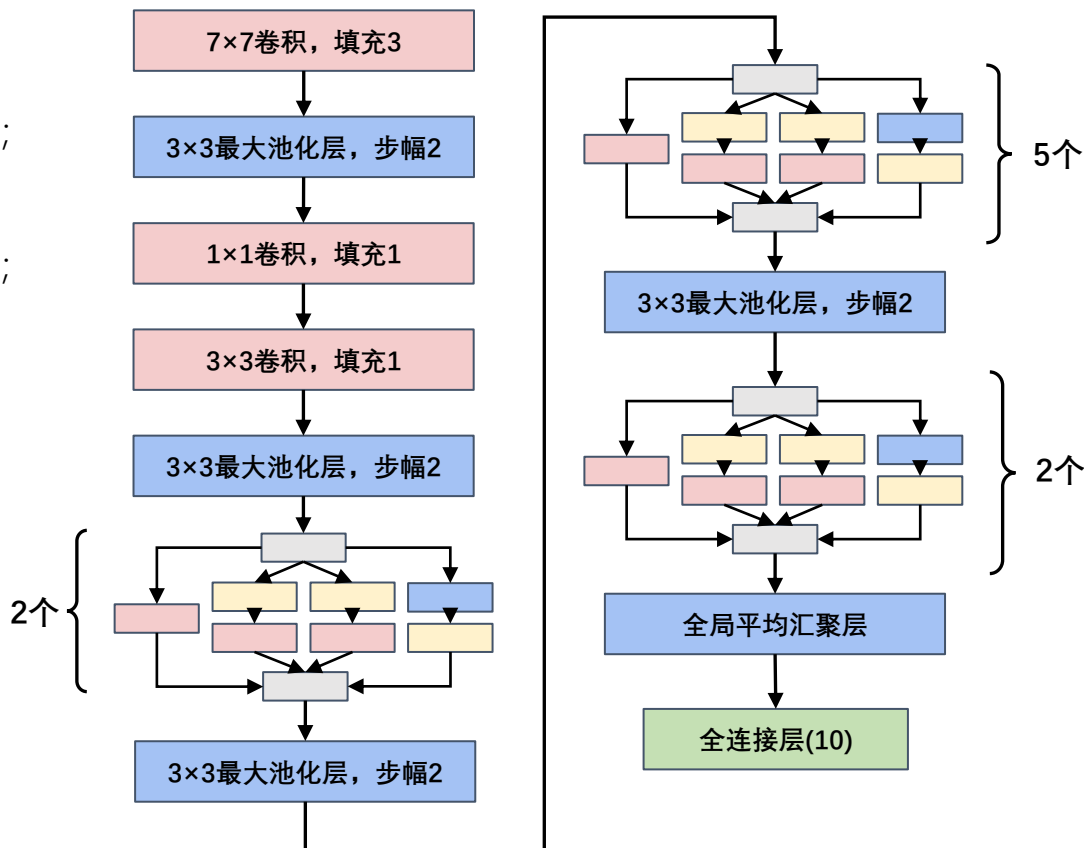
(1) 输入为 $14 \times 14 \times 512$ ，卷积核数量为24个；卷积核的尺寸大小为 $1 \times 1 \times 512$ ；步幅为1 (stride = 1)，填充为0 (padding=0)；卷积后得到shape为 $14 \times 14 \times 24$ 的特征图输出。

(2) 输入为 $14 \times 14 \times 24$ ，卷积核数量为64个；卷积核的尺寸大小为 $5 \times 5 \times 24$ ；步幅为1 (stride = 1)，填充为2 (padding=2)；卷积后得到shape为 $14 \times 14 \times 64$ 的特征图输出。

路径4:

(1) 输入为 $14 \times 14 \times 512$ ，池化核的尺寸大小为 3×3 ；步幅为1 (stride = 1)，填充为1 (padding=1)；池化后得到shape为 $14 \times 14 \times 512$ 的特征图输出。

(2) 输入为 $14 \times 14 \times 512$ ，卷积核数量为64个；卷积核的尺寸大小为 $1 \times 1 \times 3$ ；步幅为1 (stride = 1)，填充为0 (padding=0)；卷积后得到shape为 $14 \times 14 \times 64$ 的特征图输出。



GoogLeNet网络参数详解--第4个Inception块

通道合并:

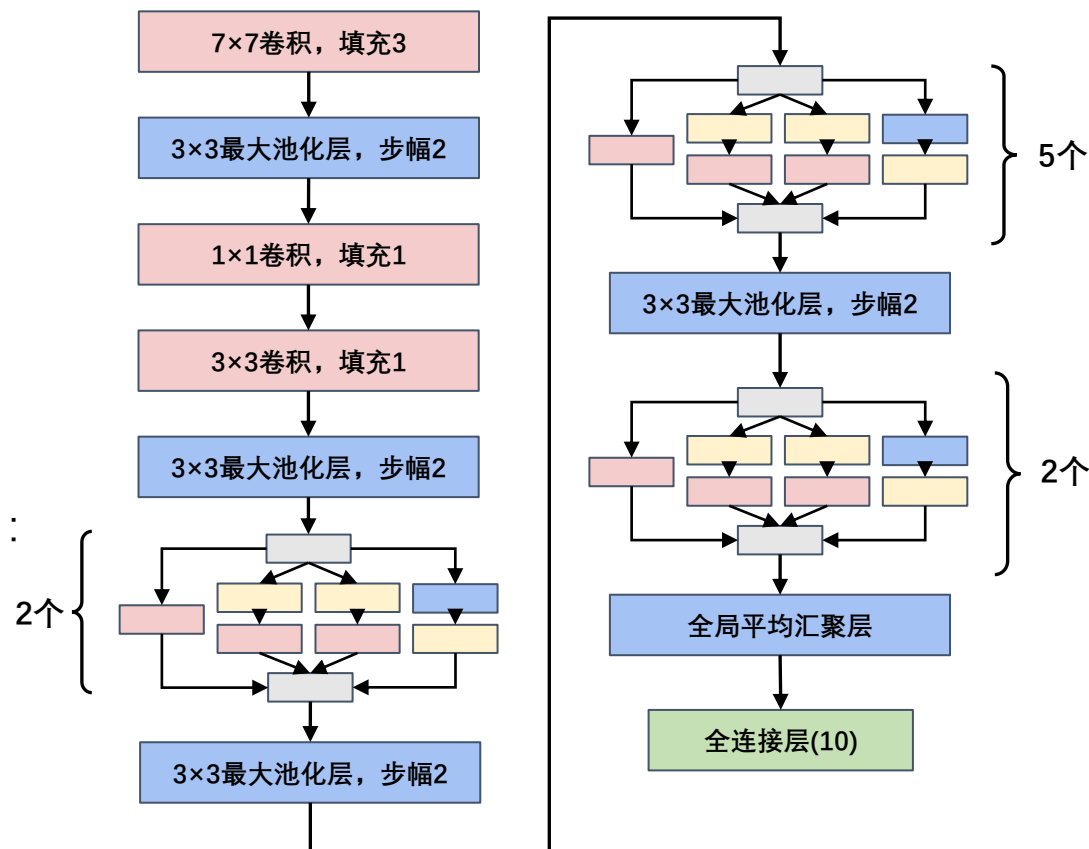
路径1的到输出为: $14 \times 14 \times 160$

路径2的到输出为: $14 \times 14 \times 224$

路径3的到输出为: $14 \times 14 \times 64$

路径4的到输出为: $14 \times 14 \times 64$

最终通道合并为 $160 + 224 + 64 + 64 = 512$, 最终的输出为:
 $14 \times 14 \times 512$ 。



GoogLeNet网络参数详解--第5个Inception块

第5个Inception块:

输入为 $14 \times 14 \times 512$ 的特征图。

路径1:

(1) 输入为 $14 \times 14 \times 512$ ，卷积核数量为128个；卷积核的尺寸大小为 $1 \times 1 \times 512$ ；步幅为1 (stride = 1)，填充为0 (padding=0)；卷积后得到shape为 $14 \times 14 \times 128$ 的特征图输出。

路径2:

(1) 输入为 $14 \times 14 \times 512$ ，卷积核数量为128个；卷积核的尺寸大小为 $1 \times 1 \times 512$ ；步幅为1 (stride = 1)，填充为0 (padding=0)；卷积后得到shape为 $14 \times 14 \times 128$ 的特征图输出。

(2) 输入为 $14 \times 14 \times 128$ ，卷积核数量为256个；卷积核的尺寸大小为 $3 \times 3 \times 128$ ；步幅为1 (stride = 1)，填充为1 (padding=1)；卷积后得到shape为 $14 \times 14 \times 256$ 的特征图输出。

路径3:

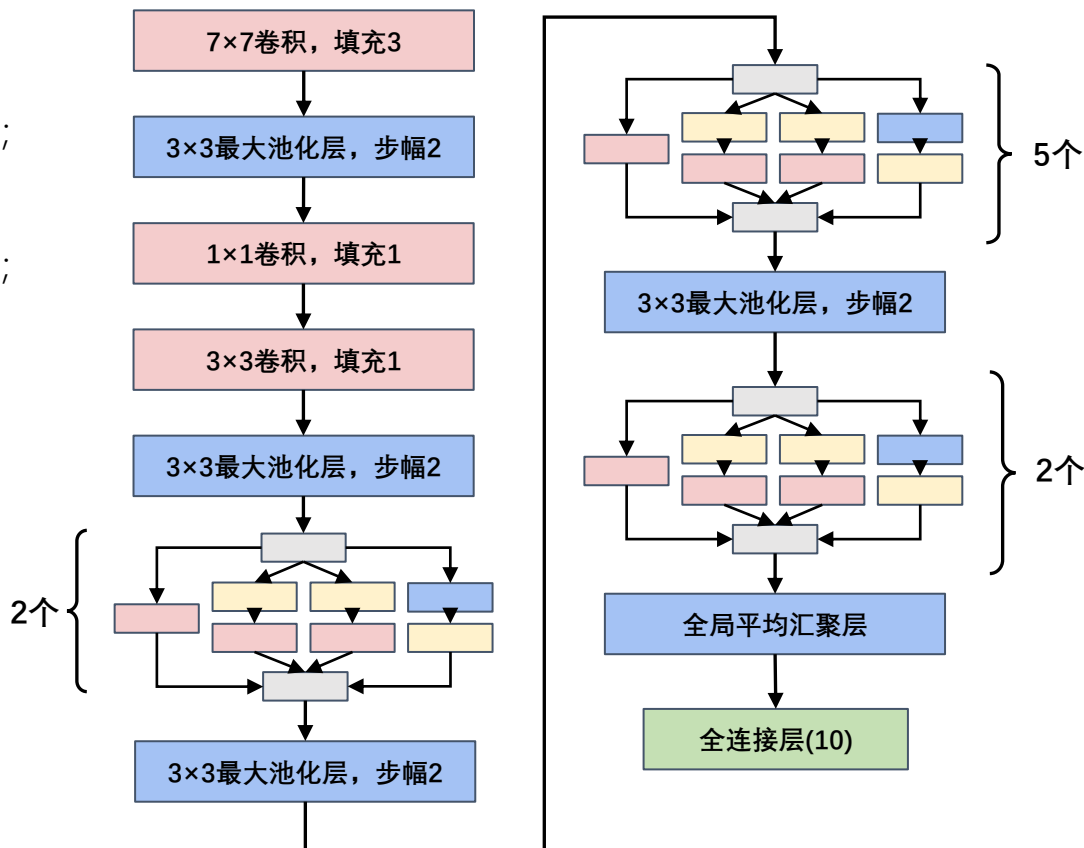
(1) 输入为 $14 \times 14 \times 512$ ，卷积核数量为24个；卷积核的尺寸大小为 $1 \times 1 \times 512$ ；步幅为1 (stride = 1)，填充为0 (padding=0)；卷积后得到shape为 $14 \times 14 \times 24$ 的特征图输出。

(2) 输入为 $14 \times 14 \times 24$ ，卷积核数量为64个；卷积核的尺寸大小为 $5 \times 5 \times 24$ ；步幅为1 (stride = 1)，填充为2 (padding=2)；卷积后得到shape为 $14 \times 14 \times 64$ 的特征图输出。

路径4:

(1) 输入为 $14 \times 14 \times 512$ ，池化核的尺寸大小为 3×3 ；步幅为1 (stride = 1)，填充为1 (padding=1)；池化后得到shape为 $14 \times 14 \times 512$ 的特征图输出。

(2) 输入为 $14 \times 14 \times 512$ ，卷积核数量为64个；卷积核的尺寸大小为 $1 \times 1 \times 3$ ；步幅为1 (stride = 1)，填充为0 (padding=0)；卷积后得到shape为 $14 \times 14 \times 64$ 的特征图输出。



GoogLeNet网络参数详解--第5个Inception块

通道合并:

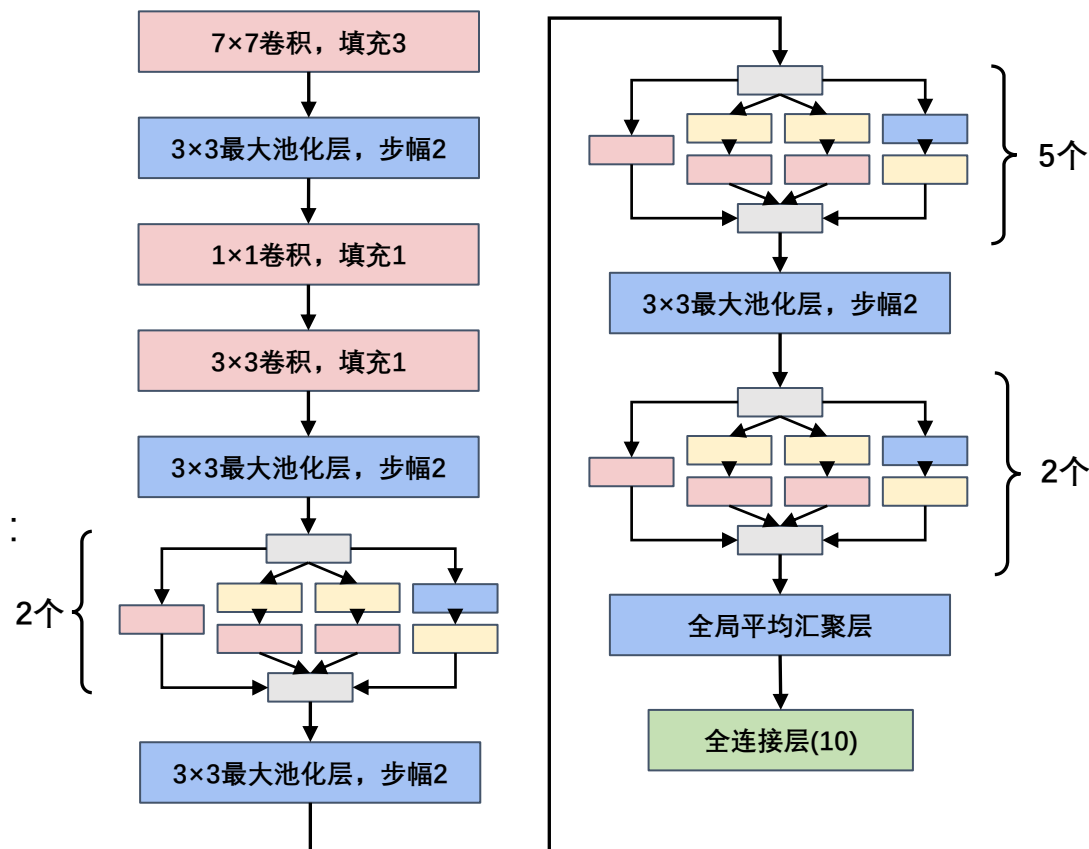
路径1的到输出为: $14 \times 14 \times 128$

路径2的到输出为: $14 \times 14 \times 256$

路径3的到输出为: $14 \times 14 \times 64$

路径4的到输出为: $14 \times 14 \times 64$

最终通道合并为 $128 + 256 + 64 + 64 = 512$, 最终的输出为:
 $14 \times 14 \times 512$ 。



GoogLeNet网络参数详解--第6个Inception块

第6个Inception块:

输入为 $14 \times 14 \times 512$ 的特征图。

路径1:

(1) 输入为 $14 \times 14 \times 512$ ，卷积核数量为112个；卷积核的尺寸大小为 $1 \times 1 \times 512$ ；步幅为1 (stride = 1)，填充为0 (padding=0)；卷积后得到shape为 $14 \times 14 \times 112$ 的特征图输出。

路径2:

(1) 输入为 $14 \times 14 \times 512$ ，卷积核数量为128个；卷积核的尺寸大小为 $1 \times 1 \times 512$ ；步幅为1 (stride = 1)，填充为0 (padding=0)；卷积后得到shape为 $14 \times 14 \times 128$ 的特征图输出。

(2) 输入为 $14 \times 14 \times 128$ ，卷积核数量为288个；卷积核的尺寸大小为 $3 \times 3 \times 128$ ；步幅为1 (stride = 1)，填充为1 (padding=1)；卷积后得到shape为 $14 \times 14 \times 288$ 的特征图输出。

路径3:

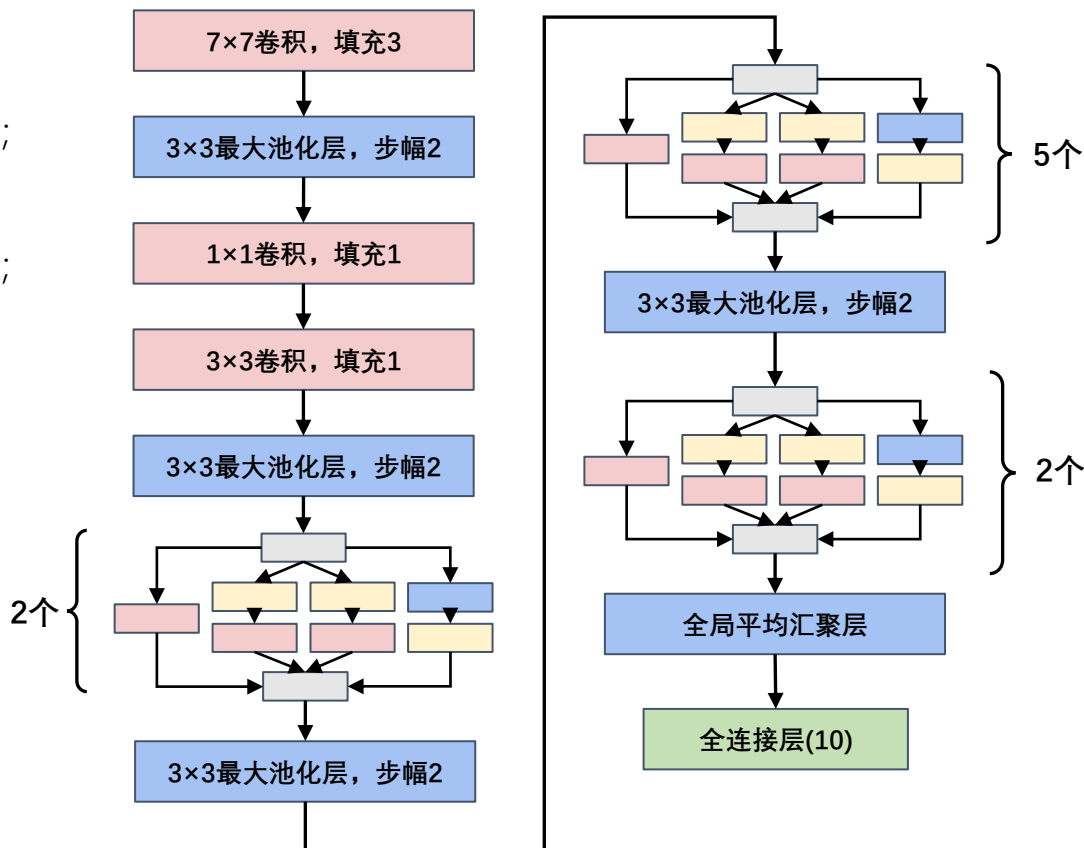
(1) 输入为 $14 \times 14 \times 512$ ，卷积核数量为32个；卷积核的尺寸大小为 $1 \times 1 \times 512$ ；步幅为1 (stride = 1)，填充为0 (padding=0)；卷积后得到shape为 $14 \times 14 \times 32$ 的特征图输出。

(2) 输入为 $14 \times 14 \times 32$ ，卷积核数量为64个；卷积核的尺寸大小为 $5 \times 5 \times 32$ ；步幅为1 (stride = 1)，填充为2 (padding=2)；卷积后得到shape为 $14 \times 14 \times 64$ 的特征图输出。

路径4:

(1) 输入为 $14 \times 14 \times 512$ ，池化核的尺寸大小为 3×3 ；步幅为1 (stride = 1)，填充为1 (padding=1)；池化后得到shape为 $14 \times 14 \times 512$ 的特征图输出。

(2) 输入为 $14 \times 14 \times 512$ ，卷积核数量为64个；卷积核的尺寸大小为 $1 \times 1 \times 3$ ；步幅为1 (stride = 1)，填充为0 (padding=0)；卷积后得到shape为 $14 \times 14 \times 64$ 的特征图输出。



GoogLeNet网络参数详解--第6个Inception块

通道合并:

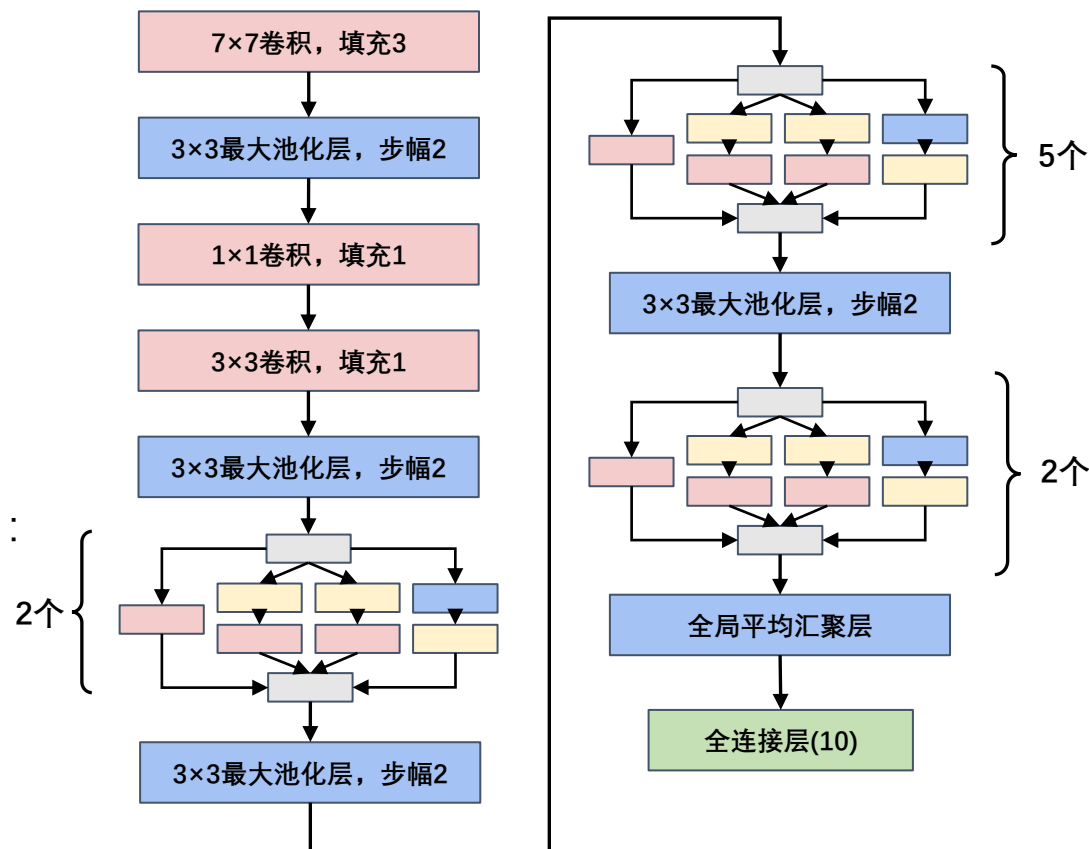
路径1的到输出为: $14 \times 14 \times 112$

路径2的到输出为: $14 \times 14 \times 288$

路径3的到输出为: $14 \times 14 \times 64$

路径4的到输出为: $14 \times 14 \times 64$

最终通道合并为 $112+288+64+64=528$, 最终的输出为:
 $14 \times 14 \times 528$ 。



GoogLeNet网络参数详解--第7个Inception块

第7个Inception块:

输入为 $14 \times 14 \times 528$ 的特征图。

路径1:

(1) 输入为 $14 \times 14 \times 528$ ，卷积核数量为256个；卷积核的尺寸大小为 $1 \times 1 \times 528$ ；步幅为1 (stride = 1)，填充为0 (padding=0)；卷积后得到shape为 $14 \times 14 \times 256$ 的特征图输出。

路径2:

(1) 输入为 $14 \times 14 \times 528$ ，卷积核数量为160个；卷积核的尺寸大小为 $1 \times 1 \times 528$ ；步幅为1 (stride = 1)，填充为0 (padding=0)；卷积后得到shape为 $14 \times 14 \times 160$ 的特征图输出。

(2) 输入为 $14 \times 14 \times 160$ ，卷积核数量为320个；卷积核的尺寸大小为 $3 \times 3 \times 160$ ；步幅为1 (stride = 1)，填充为1 (padding=1)；卷积后得到shape为 $14 \times 14 \times 320$ 的特征图输出。

路径3:

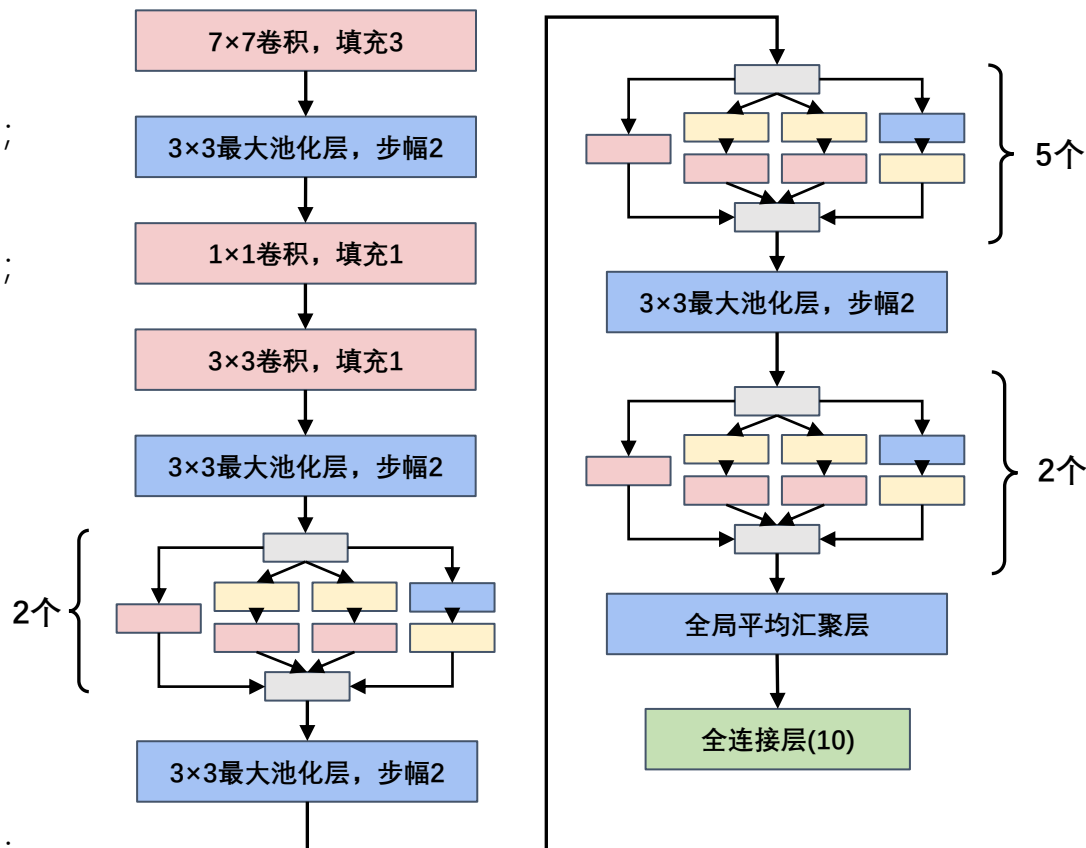
(1) 输入为 $14 \times 14 \times 528$ ，卷积核数量为32个；卷积核的尺寸大小为 $1 \times 1 \times 528$ ；步幅为1 (stride = 1)，填充为0 (padding=0)；卷积后得到shape为 $14 \times 14 \times 32$ 的特征图输出。

(2) 输入为 $14 \times 14 \times 32$ ，卷积核数量为128个；卷积核的尺寸大小为 $5 \times 5 \times 32$ ；步幅为1 (stride = 1)，填充为2 (padding=2)；卷积后得到shape为 $14 \times 14 \times 128$ 的特征图输出。

路径4:

(1) 输入为 $14 \times 14 \times 528$ ，池化核的尺寸大小为 3×3 ；步幅为1 (stride = 1)，填充为1 (padding=1)；池化后得到shape为 $14 \times 14 \times 528$ 的特征图输出。

(2) 输入为 $14 \times 14 \times 528$ ，卷积核数量为128个；卷积核的尺寸大小为 $1 \times 1 \times 528$ ；步幅为1 (stride = 1)，填充为0 (padding=0)；卷积后得到shape为 $14 \times 14 \times 128$ 的特征图输出。



GoogLeNet网络参数详解--第7个Inception块

通道合并:

路径1的到输出为: $14 \times 14 \times 256$

路径2的到输出为: $14 \times 14 \times 320$

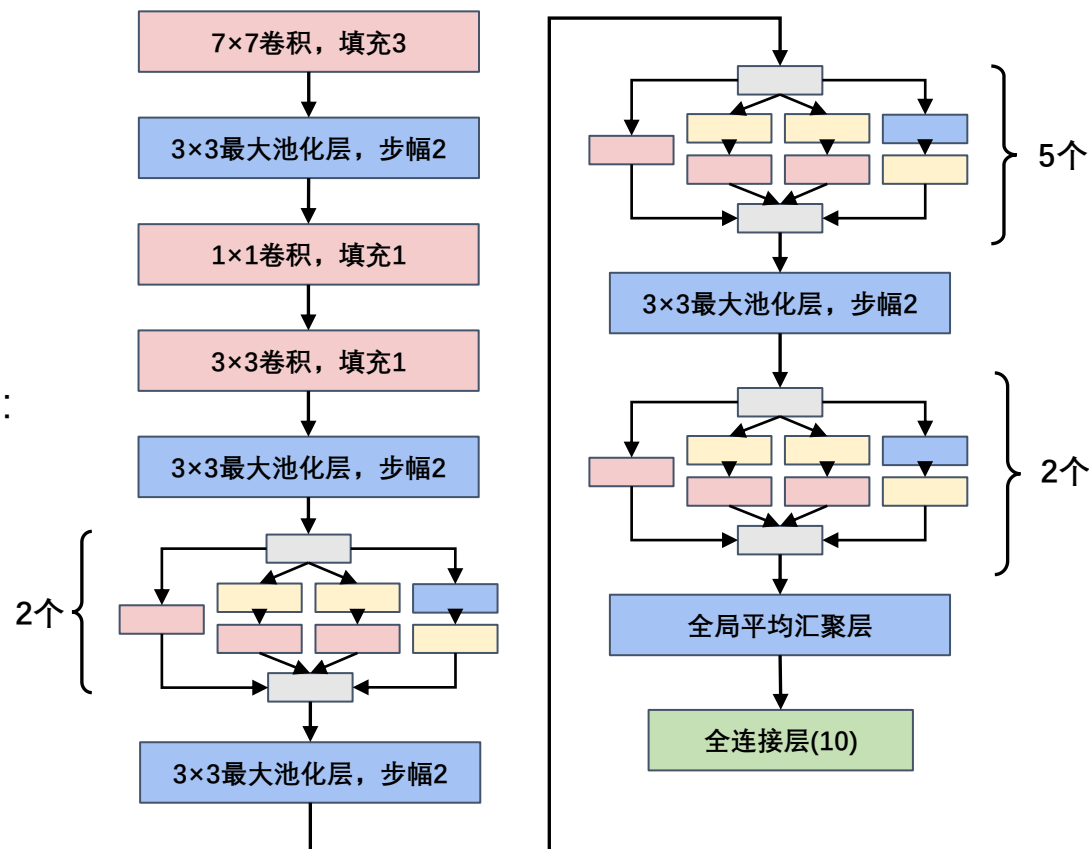
路径3的到输出为: $14 \times 14 \times 128$

路径4的到输出为: $14 \times 14 \times 128$

最终通道合并为 $256+320+128+128=832$, 最终的输出为:
 $14 \times 14 \times 832$ 。

最大池化模块:

输入为 $14 \times 14 \times 832$ 。池化核的尺寸大小为 3×3 ; 步幅为2 (stride = 2), 填充为1 (padding=1); 池化后得到shape为 $7 \times 7 \times 832$ 的特征图输出。



GoogLeNet网络参数详解--第8个Inception块

第8个Inception块:

输入为 $7 \times 7 \times 832$ 的特征图。

路径1:

(1) 输入为 $7 \times 7 \times 832$, 卷积核数量为256个; 卷积核的尺寸大小为 $1 \times 1 \times 832$; 步幅为1 (stride = 1), 填充为0 (padding=0); 卷积后得到shape为 $7 \times 7 \times 256$ 的特征图输出。

路径2:

(1) 输入为 $7 \times 7 \times 832$, 卷积核数量为160个; 卷积核的尺寸大小为 $1 \times 1 \times 832$; 步幅为1 (stride = 1), 填充为0 (padding=0); 卷积后得到shape为 $7 \times 7 \times 160$ 的特征图输出。

(2) 输入为 $7 \times 7 \times 160$, 卷积核数量为320个; 卷积核的尺寸大小为 $3 \times 3 \times 160$; 步幅为1 (stride = 1), 填充为1 (padding=1); 卷积后得到shape为 $7 \times 7 \times 320$ 的特征图输出。

路径3:

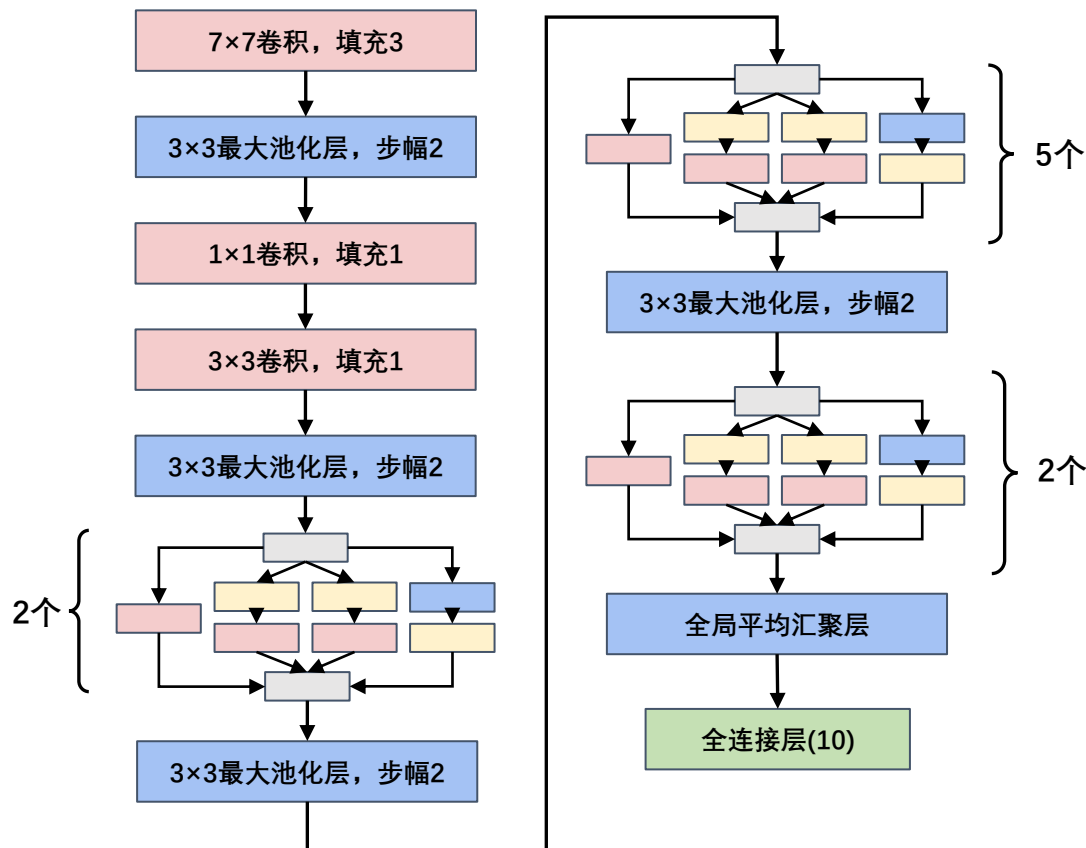
(1) 输入为 $7 \times 7 \times 832$, 卷积核数量为32个; 卷积核的尺寸大小为 $1 \times 1 \times 832$; 步幅为1 (stride = 1), 填充为0 (padding=0); 卷积后得到shape为 $7 \times 7 \times 32$ 的特征图输出。

(2) 输入为 $7 \times 7 \times 32$, 卷积核数量为128个; 卷积核的尺寸大小为 $5 \times 5 \times 32$; 步幅为1 (stride = 1), 填充为2 (padding=2); 卷积后得到shape为 $7 \times 7 \times 128$ 的特征图输出。

路径4:

(1) 输入为 $7 \times 7 \times 832$, 池化核的尺寸大小为 3×3 ; 步幅为1 (stride = 1), 填充为1 (padding=1); 池化后得到shape为 $7 \times 7 \times 832$ 的特征图输出。

(2) 输入为 $7 \times 7 \times 832$, 卷积核数量为128个; 卷积核的尺寸大小为 $1 \times 1 \times 832$; 步幅为1 (stride = 1), 填充为0 (padding=0); 卷积后得到shape为 $7 \times 7 \times 128$ 的特征图输出。



GoogLeNet网络参数详解--第8个Inception块

通道合并:

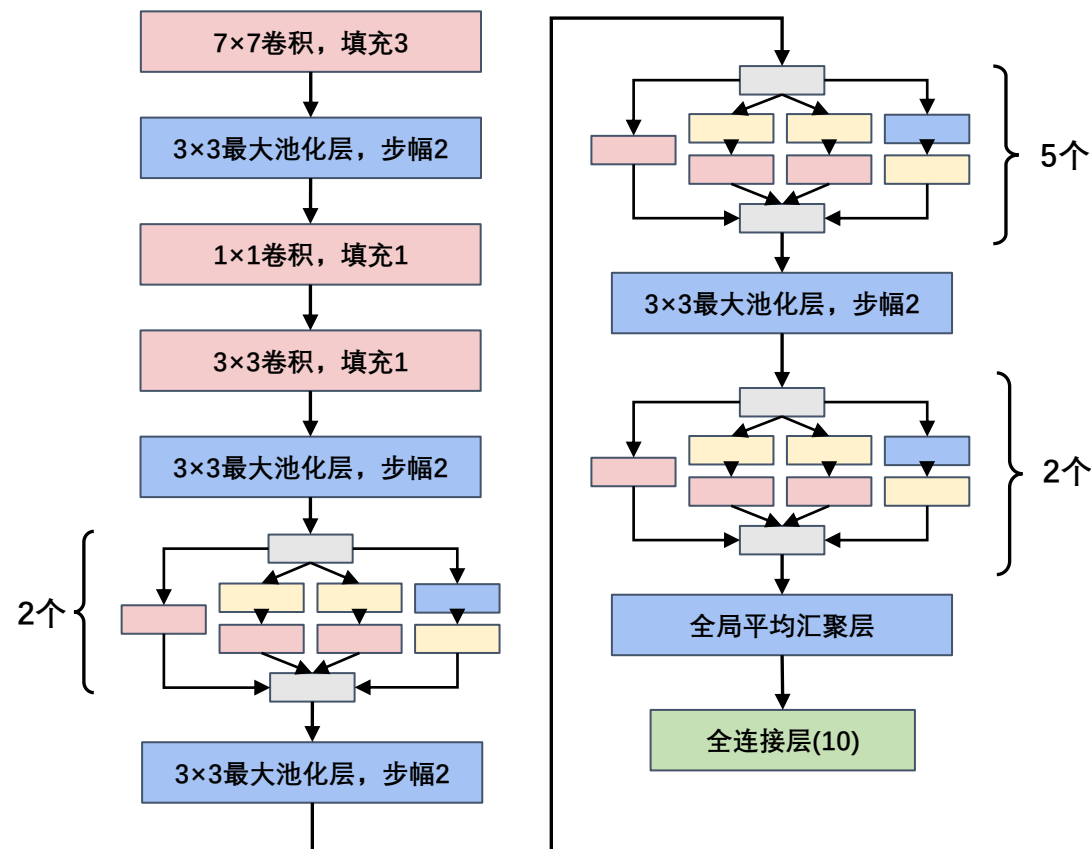
路径1的到输出为: $7 \times 7 \times 256$

路径2的到输出为: $7 \times 7 \times 320$

路径3的到输出为: $7 \times 7 \times 128$

路径4的到输出为: $7 \times 7 \times 128$

最终通道合并为 $256 + 320 + 128 + 128 = 832$, 最终的输出为:
 $7 \times 7 \times 832$ 。



GoogLeNet网络参数详解--第9个Inception块

第9个Inception块:

输入为 $7 \times 7 \times 832$ 的特征图。

路径1:

(1) 输入为 $7 \times 7 \times 832$ ，卷积核数量为384个；卷积核的尺寸大小为 $1 \times 1 \times 832$ ；步幅为1 (stride = 1)，填充为0 (padding=0)；卷积后得到shape为 $7 \times 7 \times 384$ 的特征图输出。

路径2:

(1) 输入为 $7 \times 7 \times 832$ ，卷积核数量为192个；卷积核的尺寸大小为 $1 \times 1 \times 832$ ；步幅为1 (stride = 1)，填充为0 (padding=0)；卷积后得到shape为 $7 \times 7 \times 192$ 的特征图输出。

(2) 输入为 $7 \times 7 \times 192$ ，卷积核数量为384个；卷积核的尺寸大小为 $3 \times 3 \times 192$ ；步幅为1 (stride = 1)，填充为1 (padding=1)；卷积后得到shape为 $7 \times 7 \times 384$ 的特征图输出。

路径3:

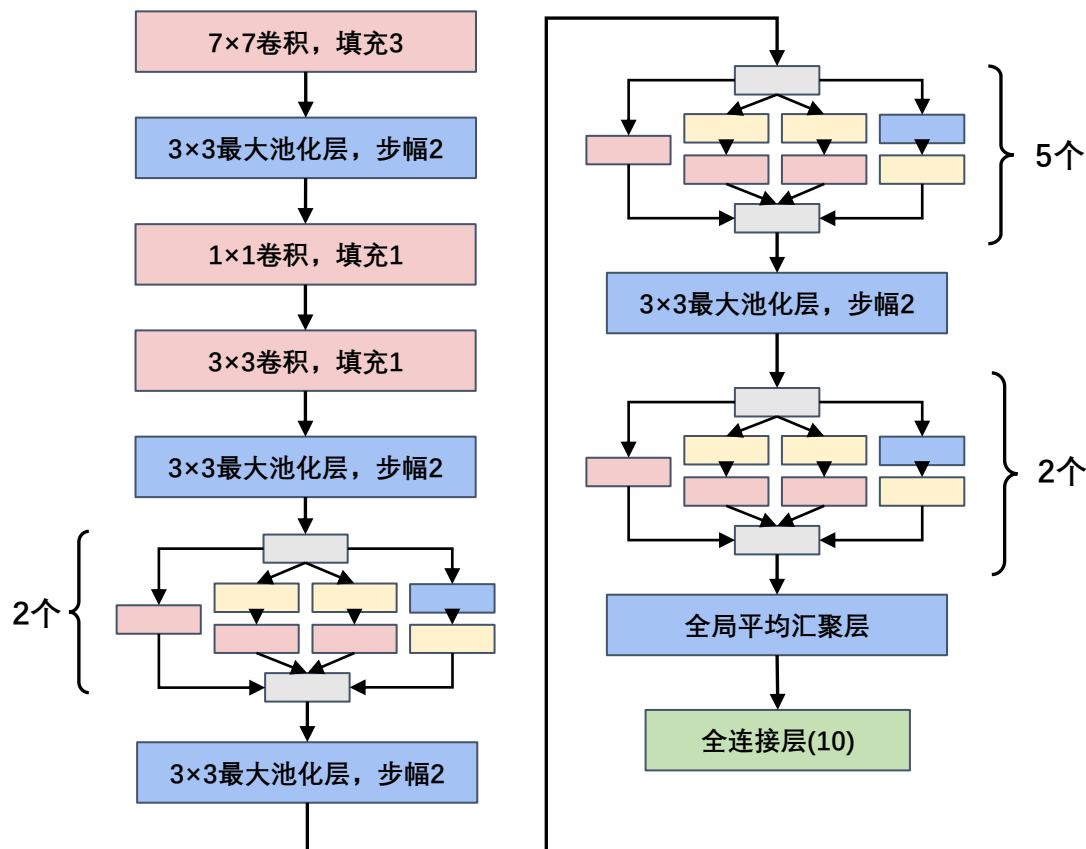
(1) 输入为 $7 \times 7 \times 832$ ，卷积核数量为48个；卷积核的尺寸大小为 $1 \times 1 \times 832$ ；步幅为1 (stride = 1)，填充为0 (padding=0)；卷积后得到shape为 $7 \times 7 \times 48$ 的特征图输出。

(2) 输入为 $7 \times 7 \times 48$ ，卷积核数量为128个；卷积核的尺寸大小为 $5 \times 5 \times 48$ ；步幅为1 (stride = 1)，填充为2 (padding=2)；卷积后得到shape为 $7 \times 7 \times 128$ 的特征图输出。

路径4:

(1) 输入为 $7 \times 7 \times 832$ ，池化核的尺寸大小为 3×3 ；步幅为1 (stride = 1)，填充为1 (padding=1)；池化后得到shape为 $7 \times 7 \times 832$ 的特征图输出。

(2) 输入为 $7 \times 7 \times 832$ ，卷积核数量为128个；卷积核的尺寸大小为 $1 \times 1 \times 832$ ；步幅为1 (stride = 1)，填充为0 (padding=0)；卷积后得到shape为 $7 \times 7 \times 128$ 的特征图输出。



GoogLeNet网络参数详解--第9个Inception块

通道合并:

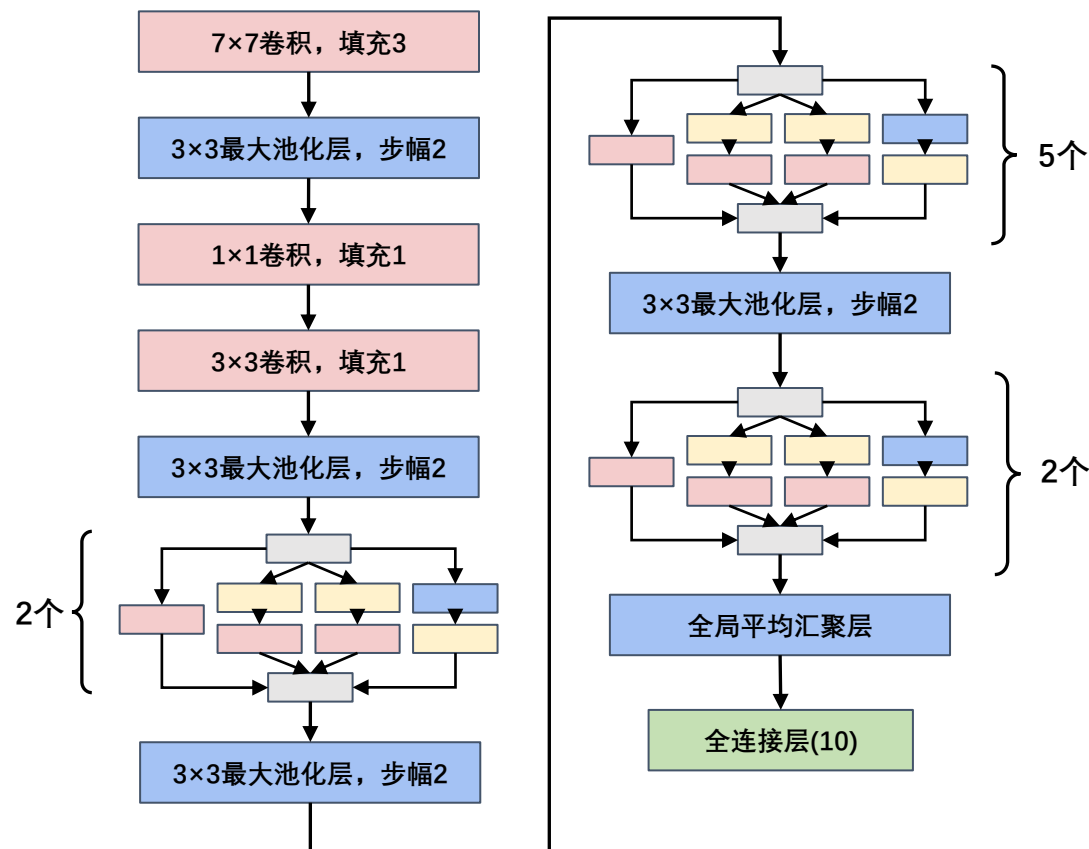
路径1的到输出为: $7 \times 7 \times 384$

路径2的到输出为: $7 \times 7 \times 384$

路径3的到输出为: $7 \times 7 \times 128$

路径4的到输出为: $7 \times 7 \times 128$

最终通道合并为 $384 + 384 + 128 + 128 = 1024$, 最终的输出为:
 $7 \times 7 \times 1024$ 。



GoogLeNet网络参数详解—输出模块

全局平均池化模块：

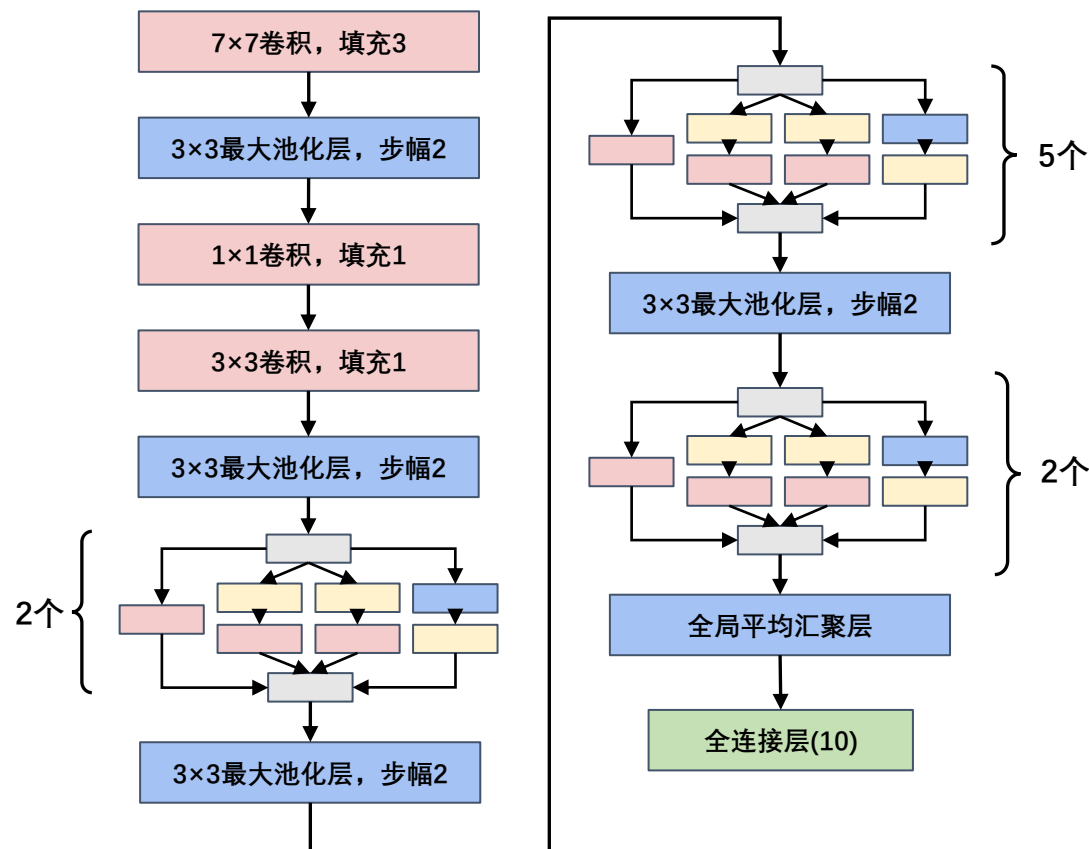
输入为 $7 \times 7 \times 1024$ 。池化后得到shape为 $1 \times 1 \times 1024$ 的特征图输出。

Flatten层：

输入为 $1 \times 1 \times 1024$ ，输出为 1×1024 。

线性全连接层：

输入为 1×1024 。线性全连接层神经元个数分别为1000。最后一层全连接层用softmax输出1000个分类。



GoogLeNet总结

GoogLeNet总结

通过创新性的Inception模块和较为复杂的网络设计，使得作者团队上力压VGGNet夺冠。而GoogLeNet作为Inception模块设计的代表，有很多值得学习的地方。通过并行的filter将输入特征提取并在通道维上进行串联合并，构成下一层的输入，再将Inception块层层叠加。通过这种方式引入了稀疏型、模拟了卷积视觉网络的最佳拓扑结构。同时也让网络具备了自动选择的能力，而不是人为地设置卷积或池化，或决定卷积核的尺寸。为了降低整体参数量，在Inception块中添加了 1×1 卷积，有效降低了特征的维度，避免了参数爆炸。

