



温州大學
WENZHOU UNIVERSITY

深度学习-深度卷积神经网络

黄海广 副教授

2022年03月

- 01** 经典网络
- 02** 深度残差网络
- 03** 谷歌**Inception** 网络
- 04** 卷积神经网络使用技巧

1.经典网络

3

01 经典网络

02 深度残差网络

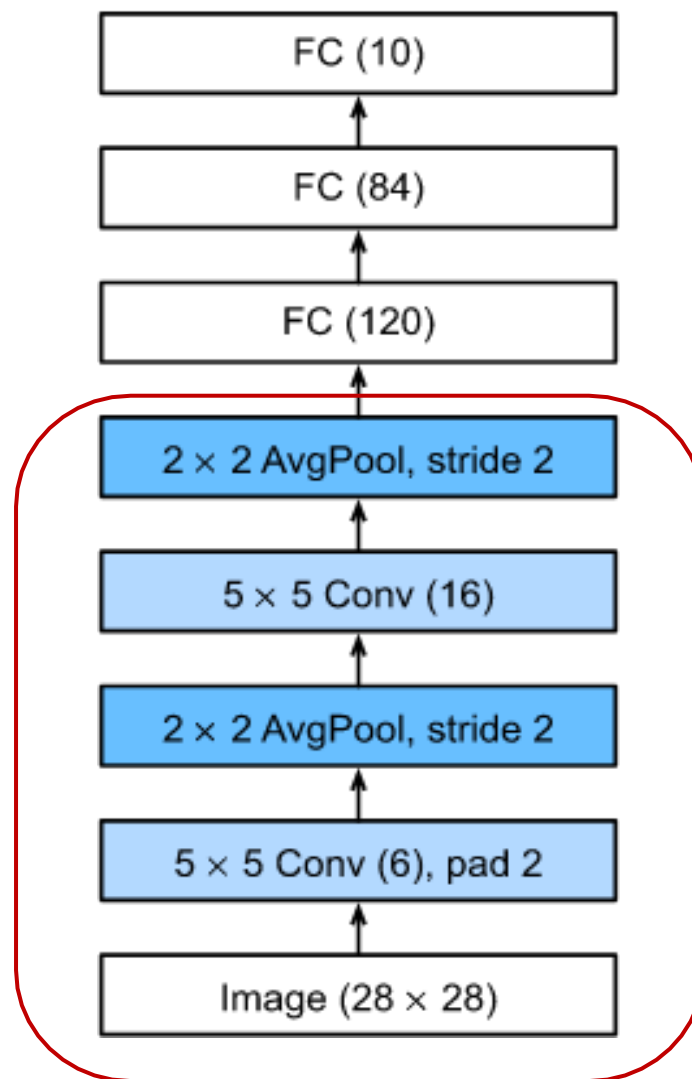
03 谷歌Inception 网络

04 卷积神经网络使用技巧

经典网络-LeNet-5

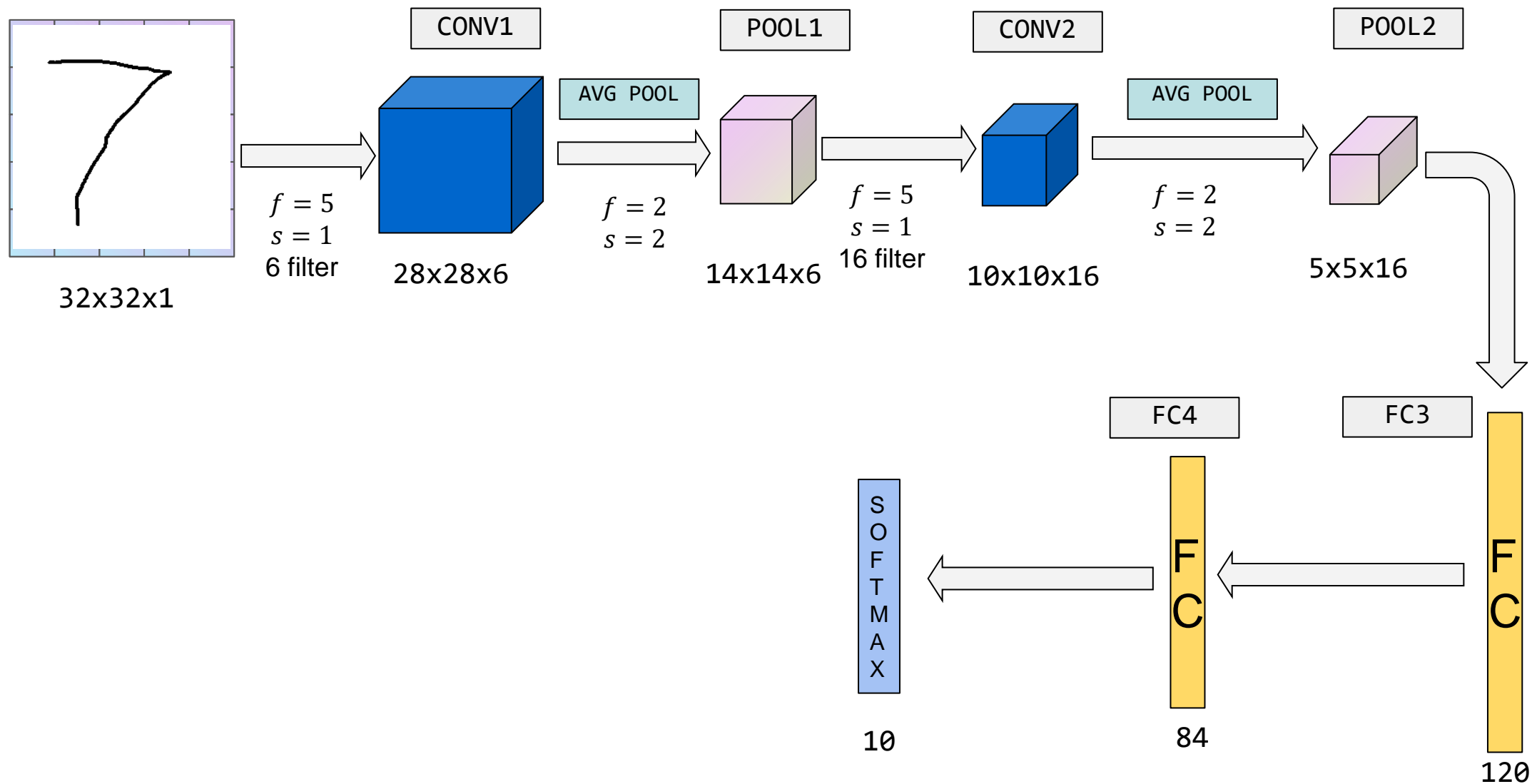
4

- LeNet 分为两个部分组成：
 - 卷积层块：由两个卷积层块组成；
 - 全连接层块：由三个全连接层组成。
- 为了将卷积层块的输出传递给全连接层块，我们必须在小批量中展平 (flatten) 每个样本。换言之，我们将这个四维输入转换成全连接层所期望的二维输入。这里的二维表示的第一个维度索引小批量中的样本，第二个维度给出每个样本的平面向量表示。LeNet 有三个全连接层，分别有 120、84 和 10 个输出。因为我们仍在执行分类，所以输出层的 10 维对应于最后输出结果的数量。



经典网络-LeNet-5

5



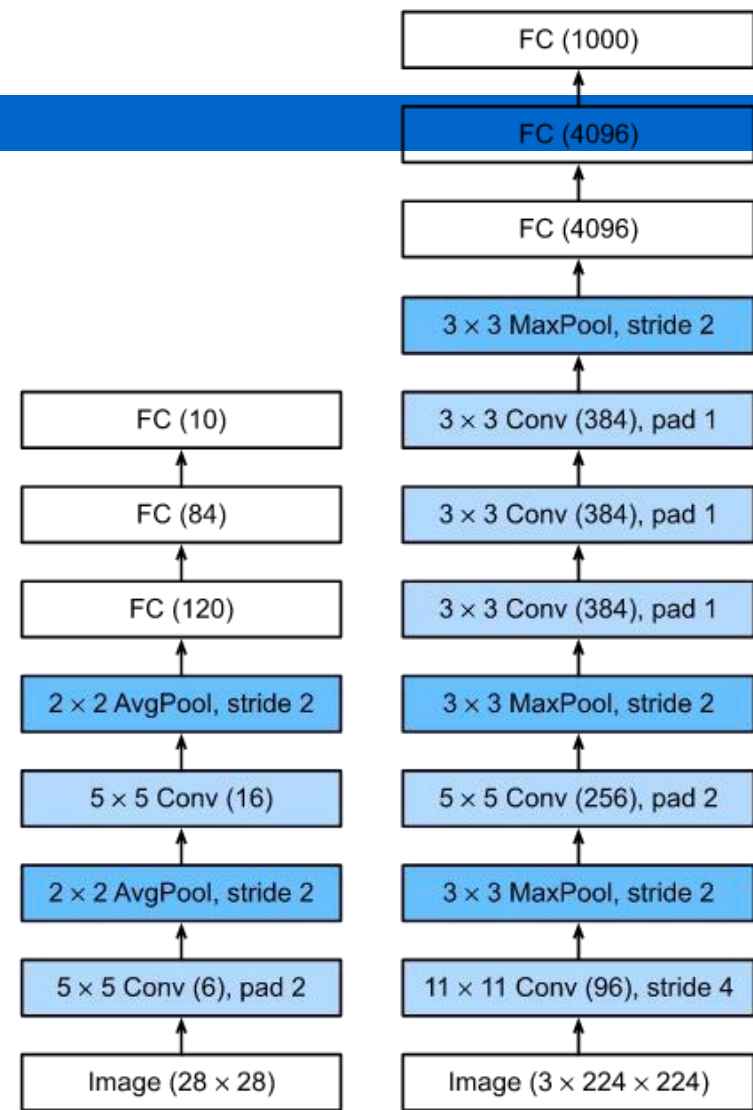
深度卷积神经网络(AlexNet)

- 2012年, AlexNet 横空出世。它首次证明了学习到的特征可以超越手工设计的特征。它一举了打破计算机视觉研究的现状。 AlexNet 使用了8层卷积神经网络, 并以很大的优势赢得了2012年 ImageNet 图像识别挑战赛。
- AlexNet 和 LeNet 的架构非常相似, 如右图所示。注意, 这里我们提供了一个稍微精简版本的 AlexNet, 去除了当年需要两个小型 GPU 同时运算的设计特点。
- AlexNet 和 LeNet 的设计理念非常相似, 但也存在显著差异。首先, AlexNet 比相对较小的 LeNet5 要深得多。 AlexNet 由八层组成: 五个卷积层、两个全连接隐藏层和一个全连接输出层。其次, AlexNet 使用 ReLU 而不是 sigmoid 作为其激活函数。
- 下面, 让我们深入研究 AlexNet 的细节。



深度卷积神经网络(AlexNet)

- 在 AlexNet 的第一层，卷积窗口的形状是 11×11 。由于大多数 ImageNet 中图像的宽和高比 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、第二层和第五层之后，加入窗口形状为 3×3 、步幅为 2 的最大池化层。此外，AlexNet 的卷积通道是 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。这两个巨大的全连接层拥有将近 1GB 的模型参数。由于早期 GPU 显存有限，原版的 AlexNet 采用了双数据流设计，使得每个 GPU 只负责存储和计算模型的一半参数。幸运的是，现在GPU显存相对充裕，所以我们现在很少需要跨 GPU 分解模型 (因此，我们的AlexNet模型在这方面与原始论文稍有不同)。

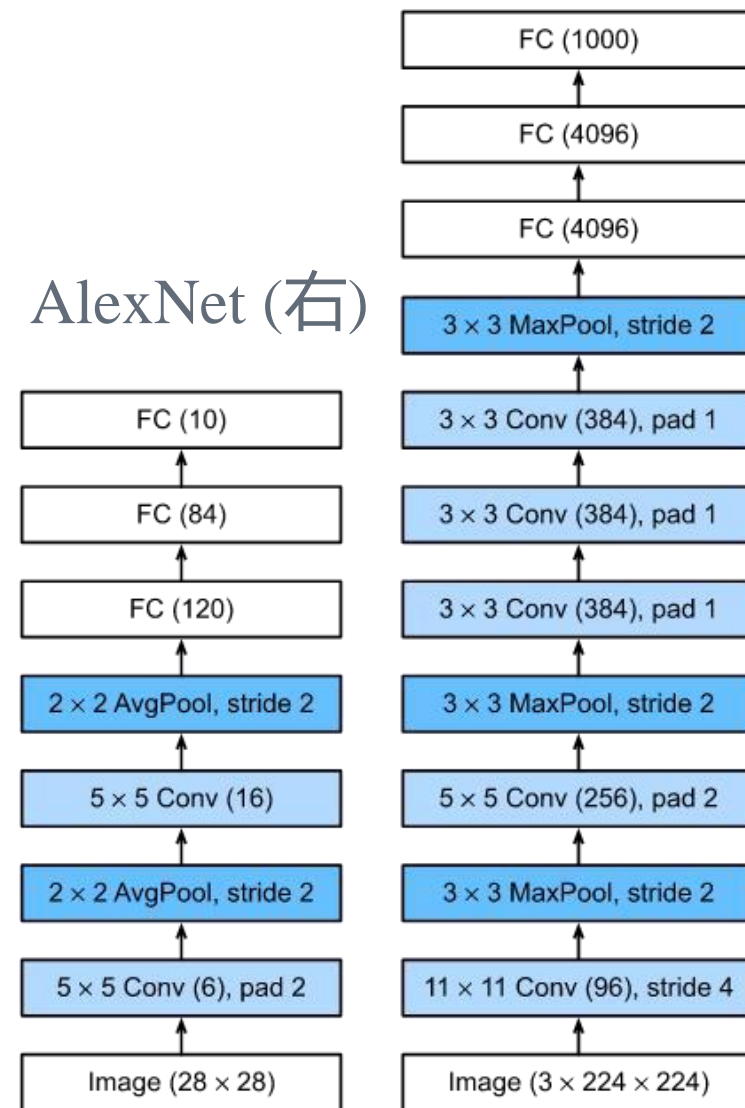


LeNet (左), AlexNet (右)

深度卷积神经网络(AlexNet)

- 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。当 sigmoid 激活函数的输出非常接近于 0 或 1 时, 这些区域的梯度几乎为 0, 因此反向传播无法继续更新一些模型参数。相反, ReLU 激活函数在正区间的梯度总是 1。
- 因此, 如果模型参数没有正确初始化, sigmoid 函数可能在正区间内得到几乎为 0 的梯度, 从而使模型无法得到有效的训练。

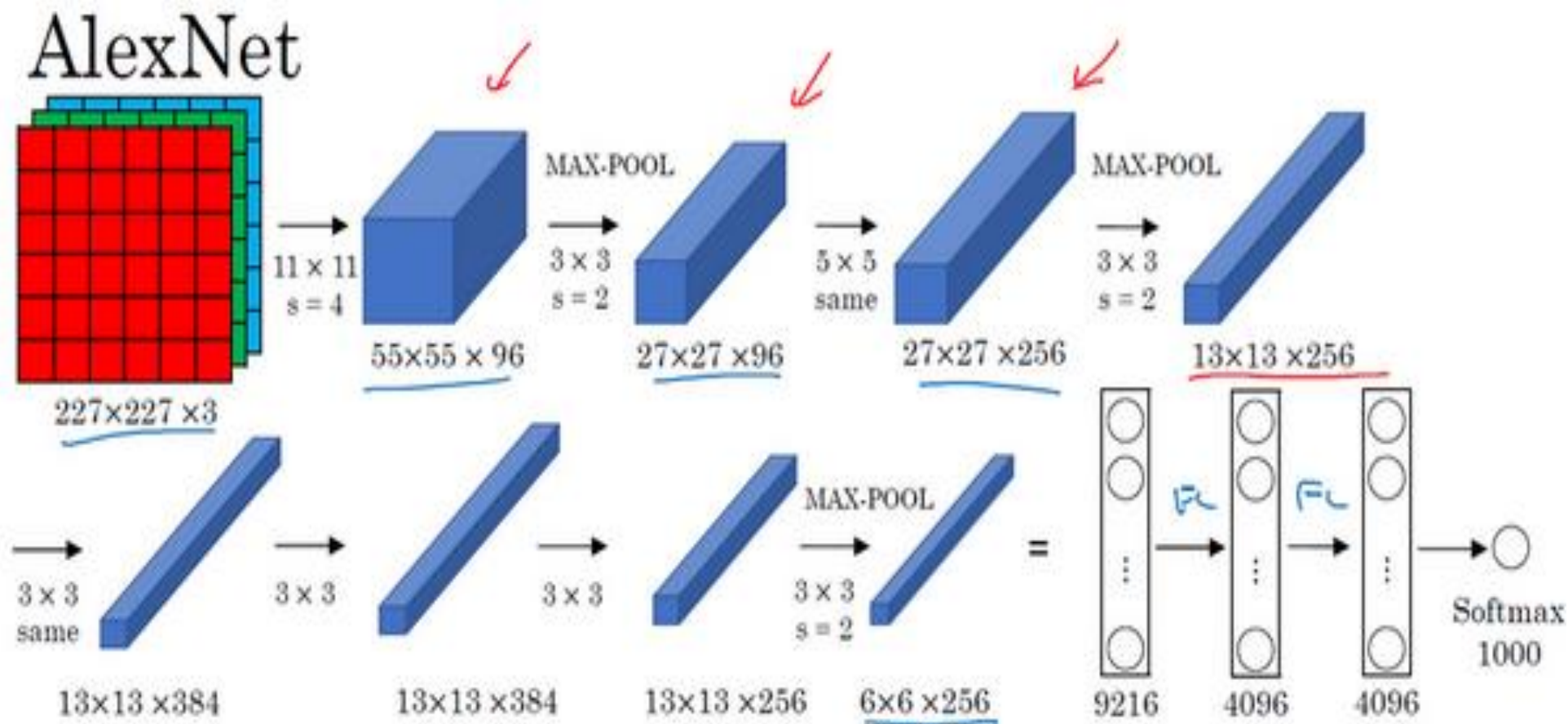
LeNet (左), AlexNet (右)



“Talk is cheap. Show me the code.”

经典网络

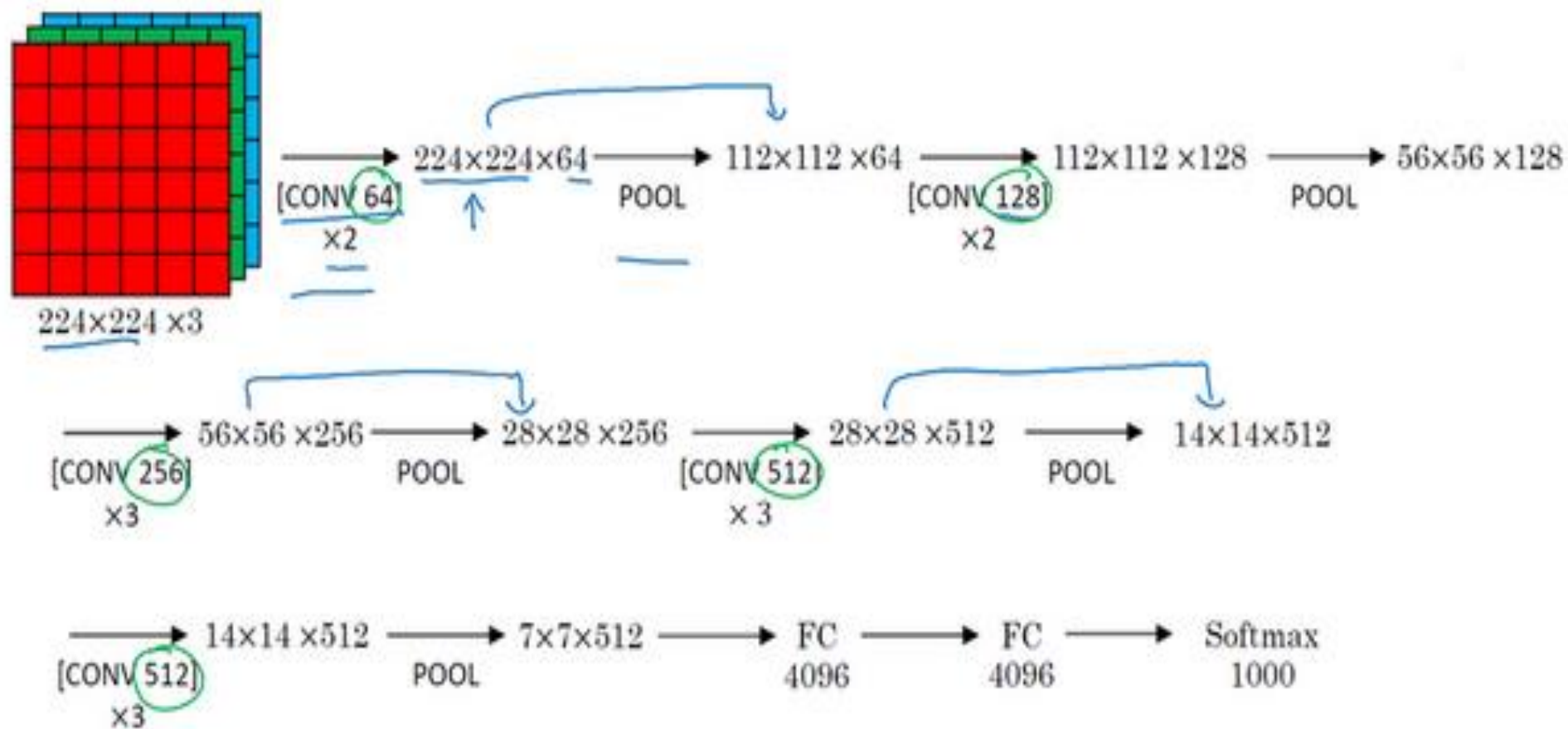
9



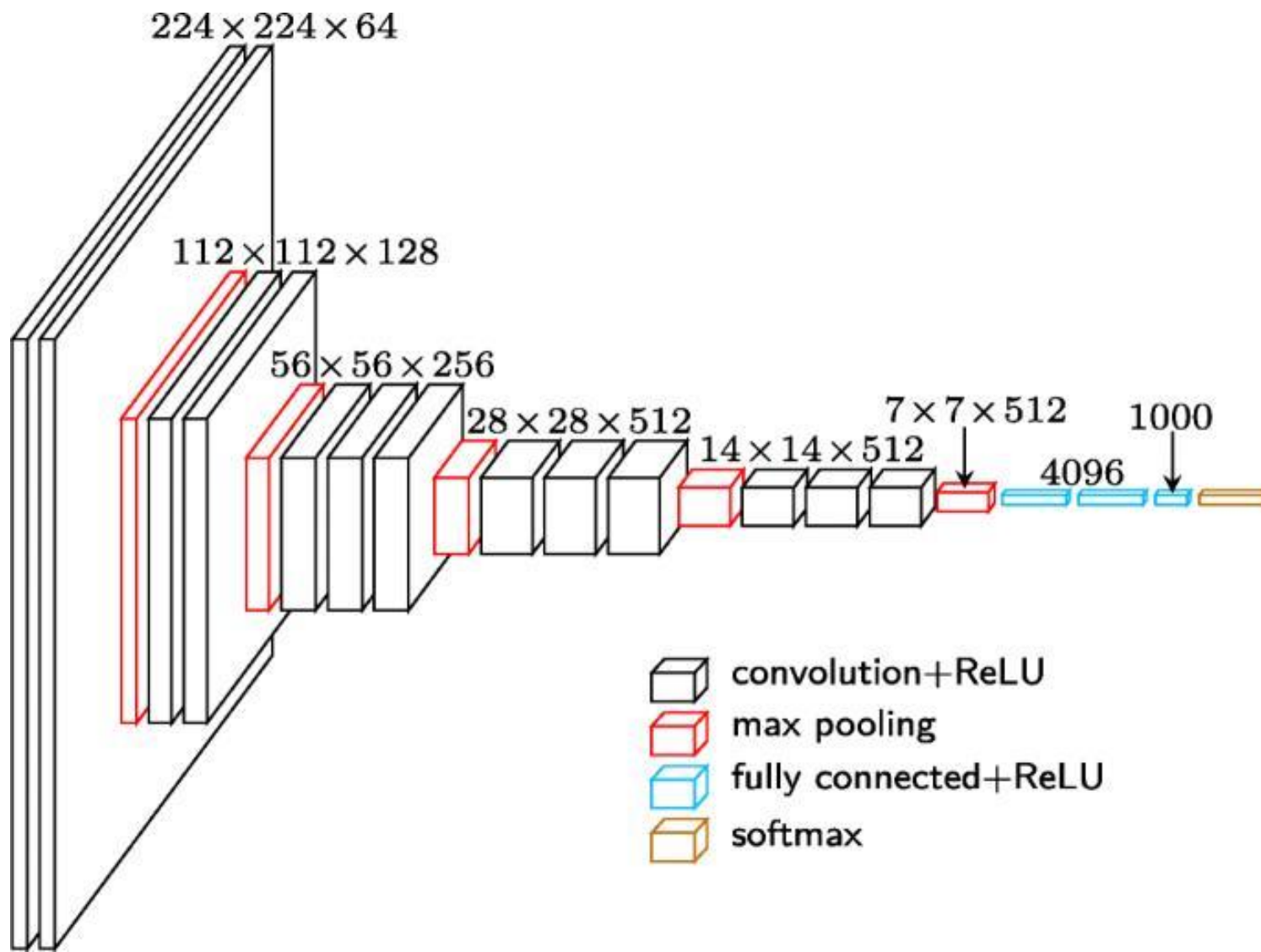
经典网络

10

VGG16



VGG16



2.深度残差网络

12

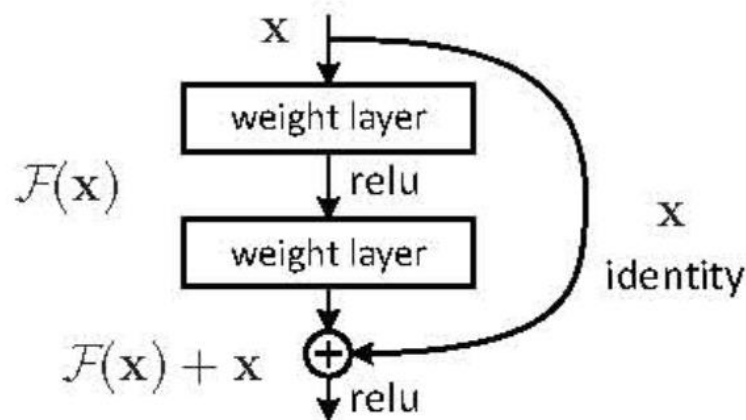
01 经典网络

02 深度残差网络

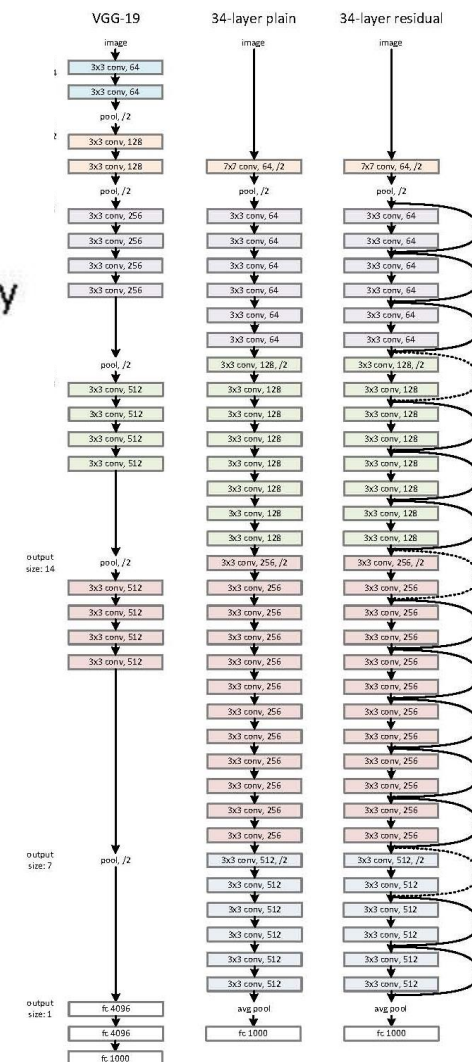
03 谷歌**Inception** 网络

04 卷积神经网络使用技巧

13

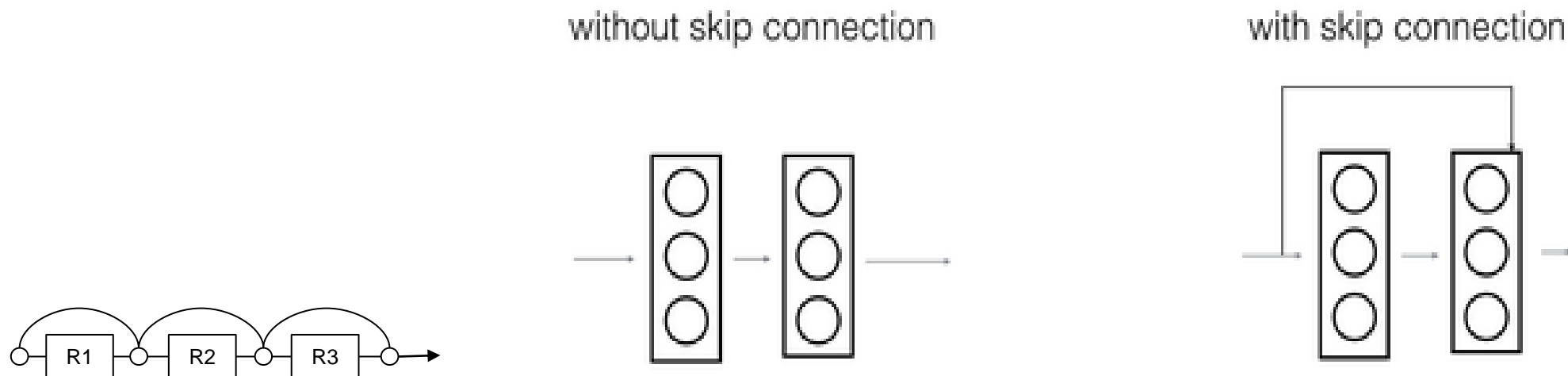


The diagram illustrates the difference between a standard function $F(x)$ and a residual connection. On the left, a gray box labeled 'Previous input' outputs x to a pink box labeled 'Stacked layers' (representing $F(x)$), which outputs $y = F(x)$. On the right, the same setup is shown, but the output of the 'Stacked layers' is added to the original input x at a summation node (a circle with a '+'). The output of the summation node is $y = F(x) + x$. A label 'x identity' points to the input x entering the summation node.

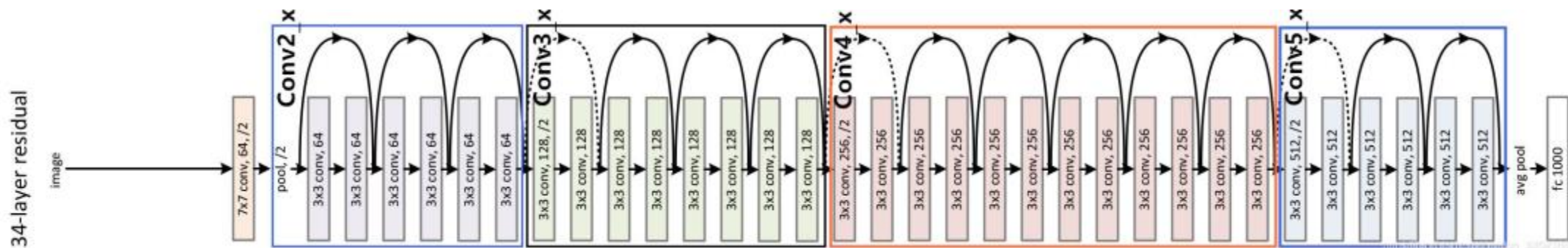


2.深度残差网络

14



ResNets使用了许多same卷积



3.谷歌Inception网络

15

01 经典网络

02 深度残差网络

03 谷歌Inception 网络

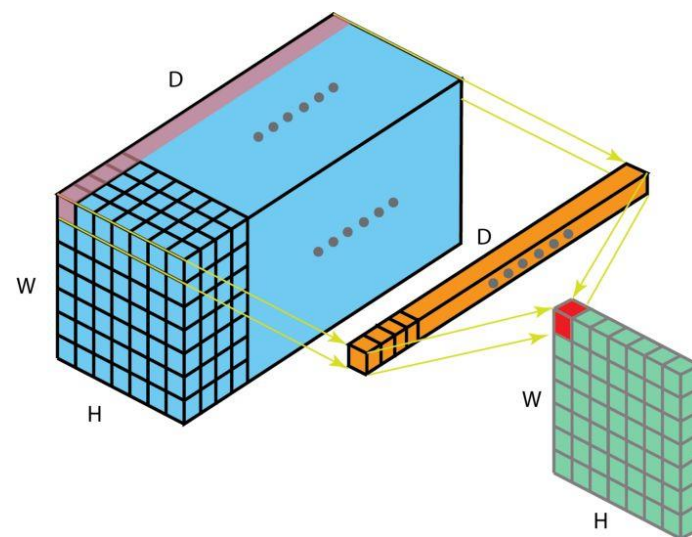
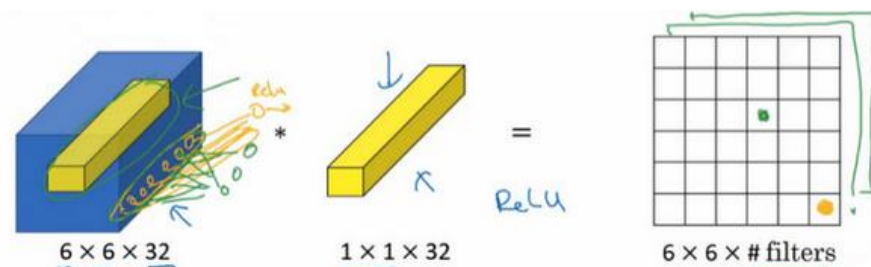
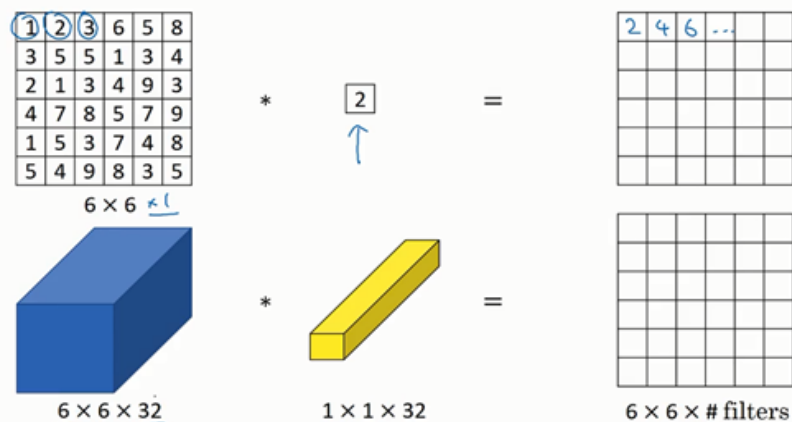
04 卷积神经网络使用技巧

3.谷歌Inception网络

16

1×1 卷积 (Network in Network)

Why does a 1×1 convolution do?

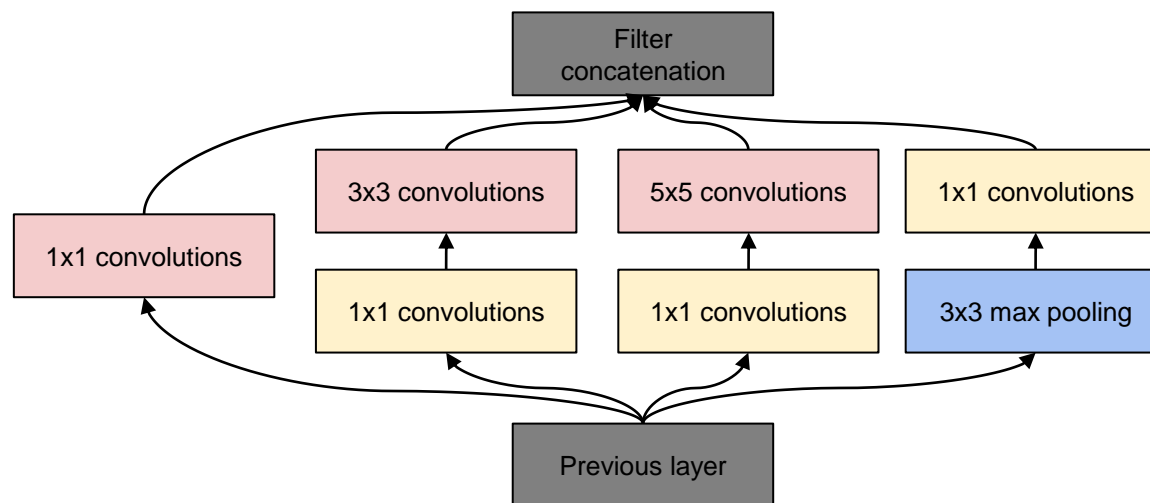
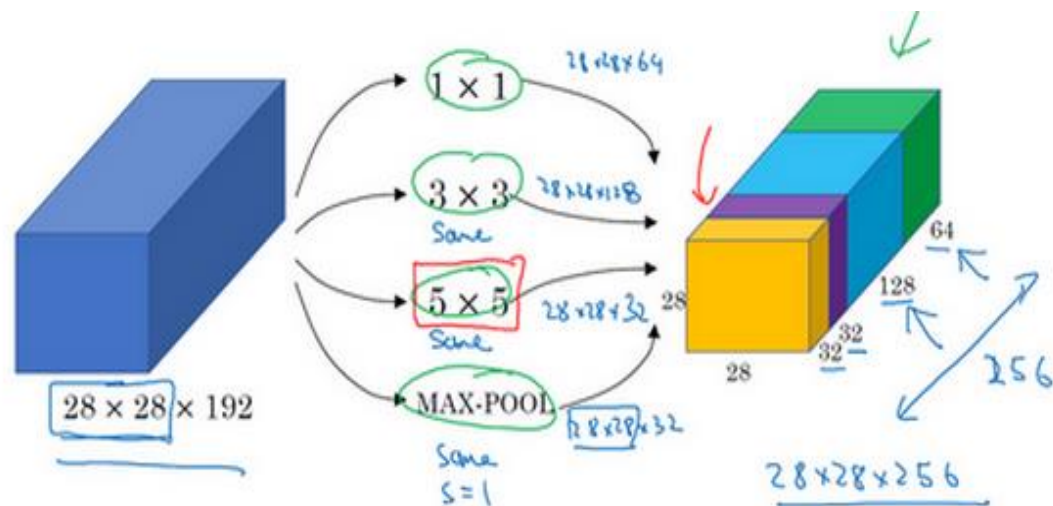


1×1 卷积层就是这样实现了一些重要功能的 (**doing something pretty non-trivial**)，它给神经网络添加了一个非线性函数，从而减少或保持输入层中的通道数量不变，当然如果你愿意，也可以增加通道数量。

3.谷歌Inception网络

17

Inception模块

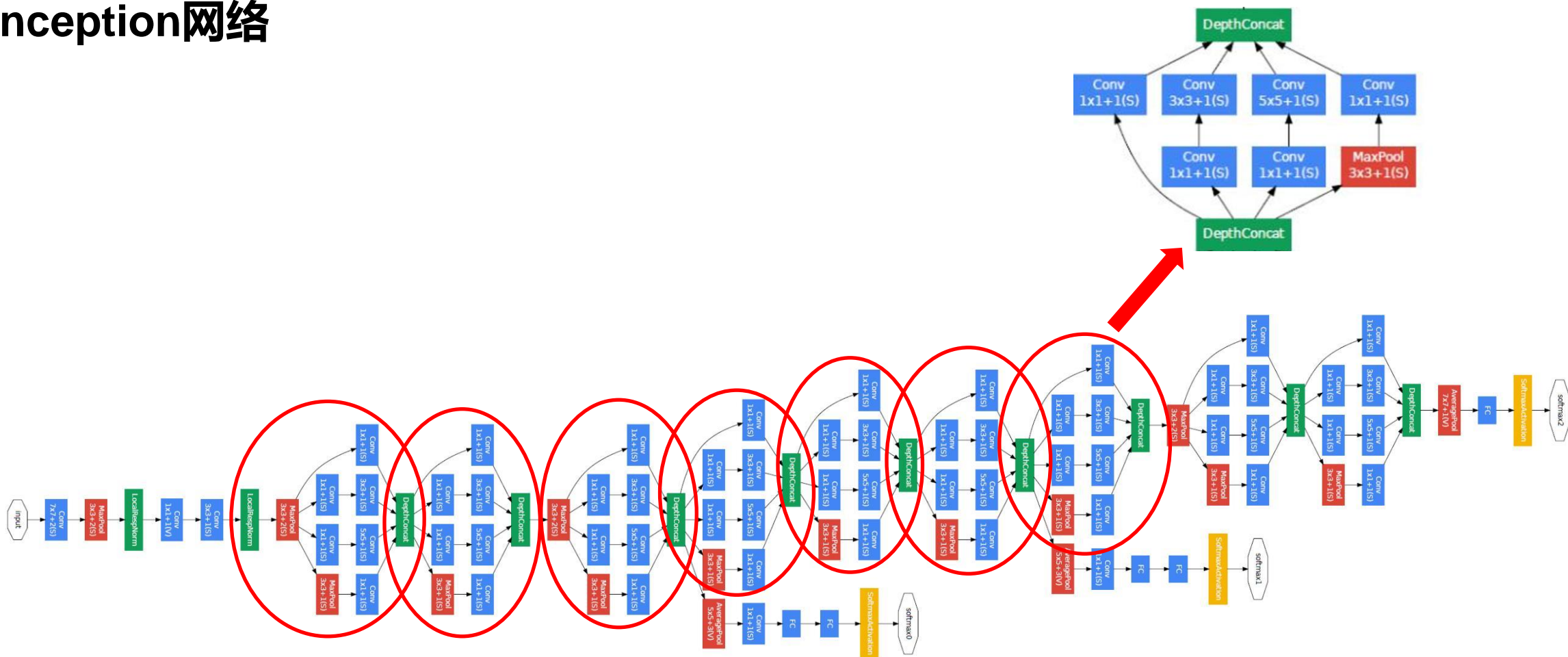


有了这样的**Inception**模块，你就可以输入某个量，因为它累加了所有数字，这里的最终输出为 $32+32+128+64=256$ 。**Inception**模块的输入为 $28 \times 28 \times 192$ ，输出为 $28 \times 28 \times 256$ 。

3.谷歌Inception网络

18

Inception网络



4.卷积神经网络使用技巧

19

01 经典网络

02 深度残差网络

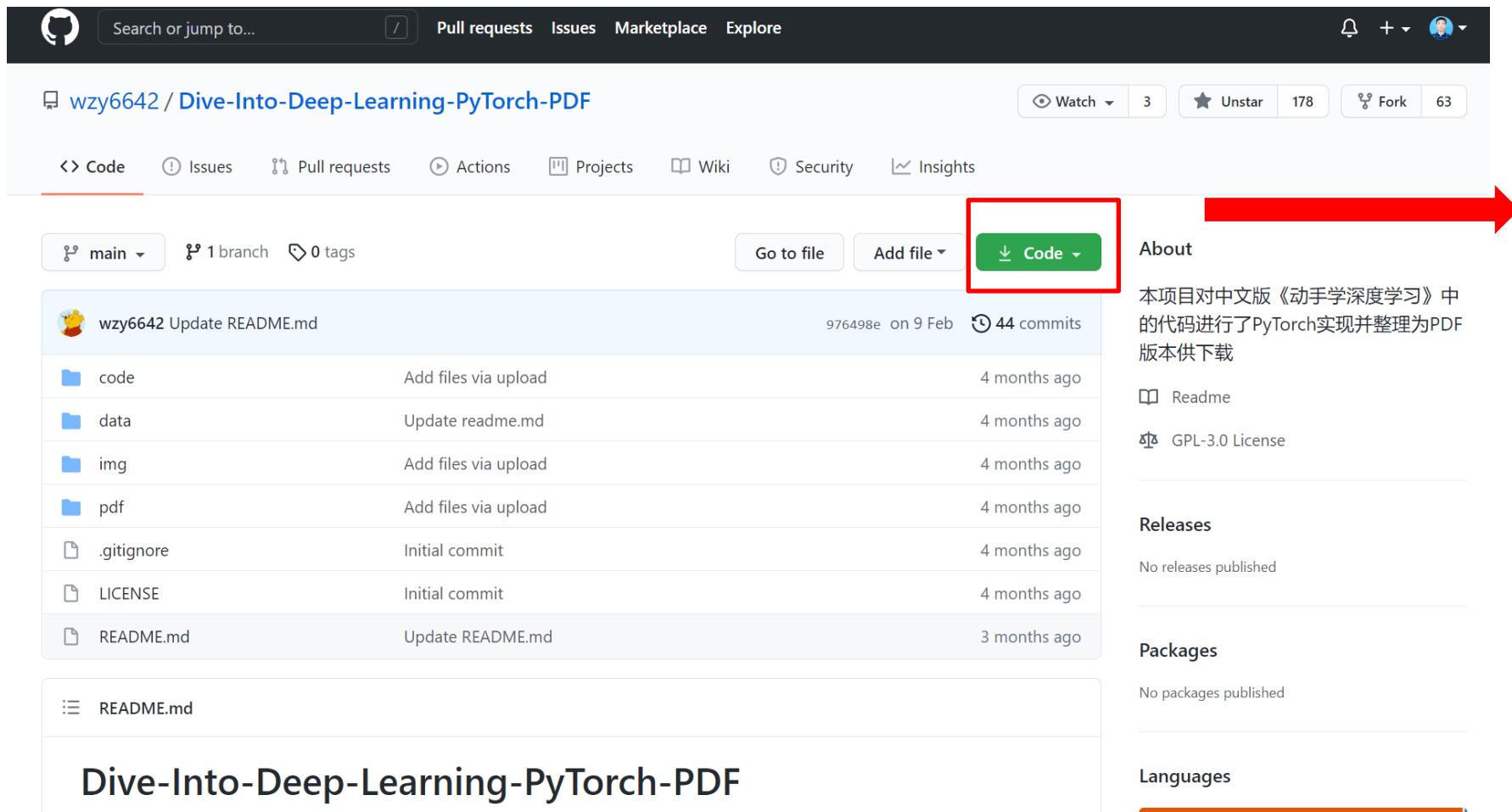
03 谷歌**Inception** 网络

04 卷积神经网络使用技巧

4.卷积神经网络使用技巧

20

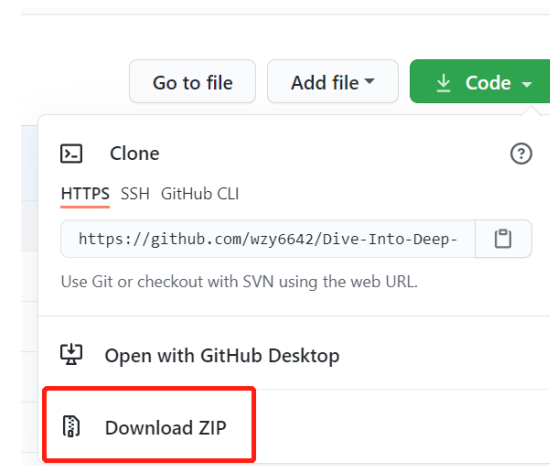
使用开源的方案



GitHub repository page for `wzy6642 / Dive-Into-Deep-Learning-PyTorch-PDF`. The page shows the repository structure with folders like `code`, `data`, `img`, and `pdf`, and files like `.gitignore`, `LICENSE`, and `README.md`. A red box highlights the `Code` button, and a red arrow points to the right, indicating the next step in the process.

File/Folder	Commit Message	Commit Hash	Commit Date	Commits
<code>code</code>	Add files via upload	976498e	on 9 Feb	44 commits
<code>data</code>	Update readme.md			4 months ago
<code>img</code>	Add files via upload			4 months ago
<code>pdf</code>	Add files via upload			4 months ago
<code>.gitignore</code>	Initial commit			4 months ago
<code>LICENSE</code>	Initial commit			4 months ago
<code>README.md</code>	Update README.md			3 months ago

Repository Name: Dive-Into-Deep-Learning-PyTorch-PDF



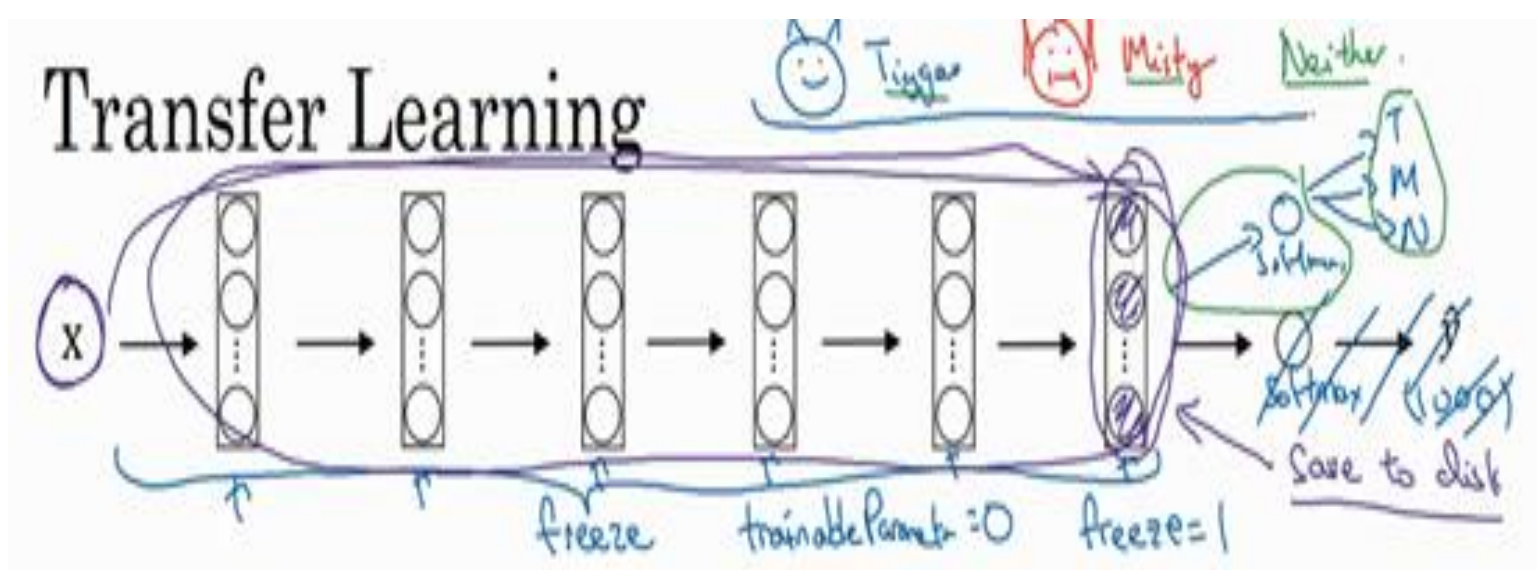
GitHub 'Code' dropdown menu options:

- Go to file
- Add file
- Code (selected)
- Clone (HTTPS, SSH, GitHub CLI)
- Open with GitHub Desktop
- Download ZIP (highlighted)

4.卷积神经网络使用技巧

21

迁移学习

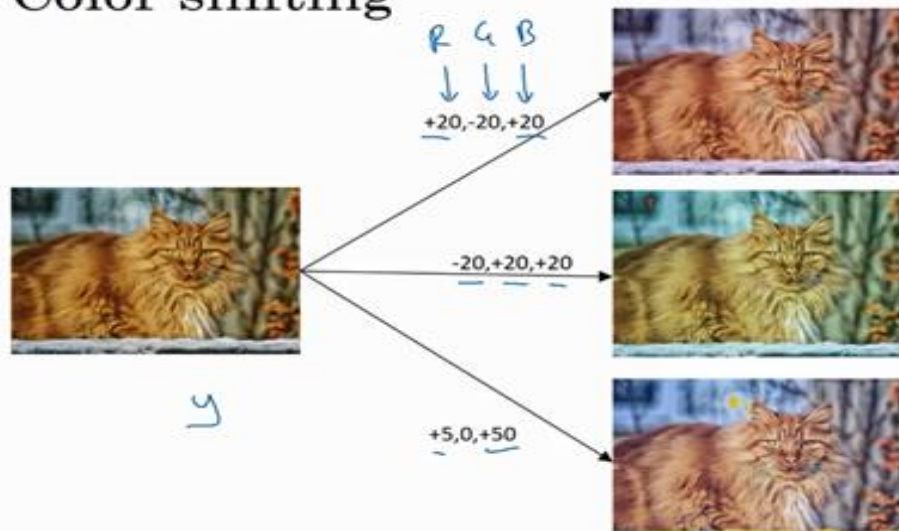


4.卷积神经网络使用技巧

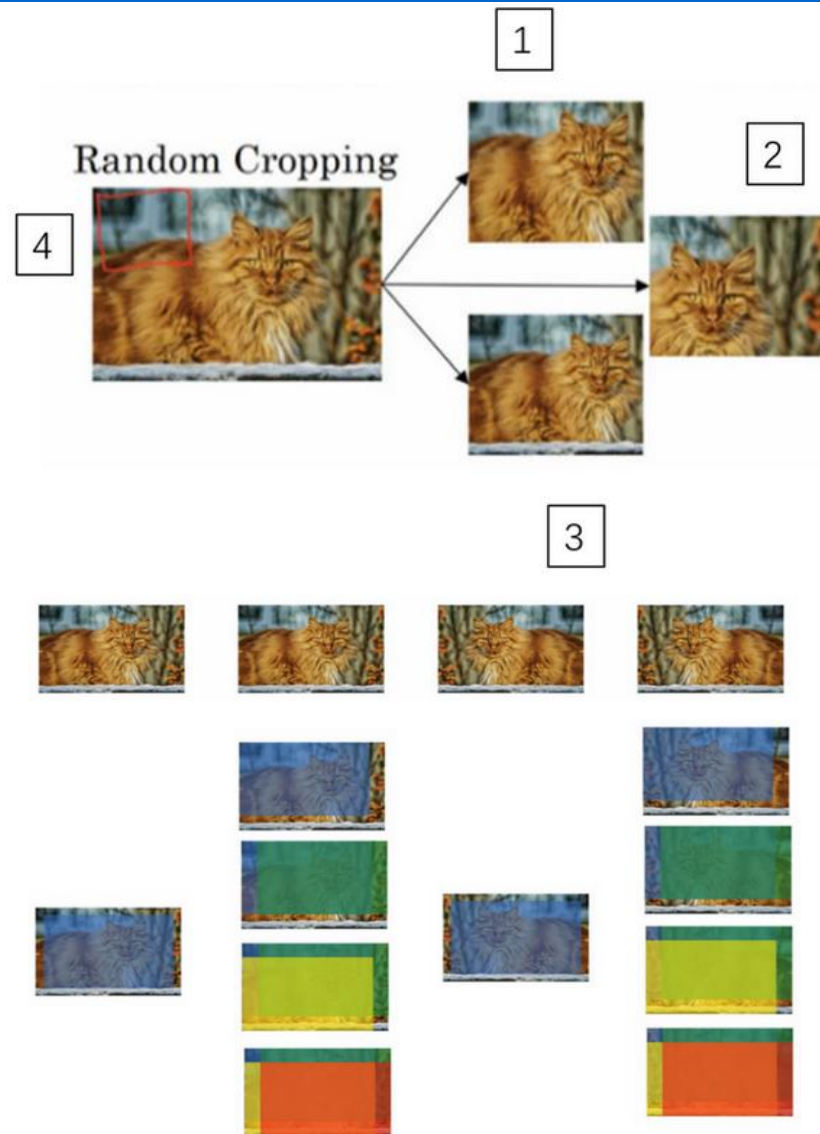
22

数据增强

Color shifting



Advanced:
PCA
ml-class.org
AlexNet paper
"PCA color
augmentation."
R B G



1. IAN GOODFELLOW等, 《深度学习》, 人民邮电出版社, 2017
2. Andrew Ng, <http://www.deeplearning.ai>

谢谢!

