

Pytorch

深度学习入门与实战

讲师：日月光华



使用预训练网络 (迁移学习)

讲师：日月光华 讲师QQ：984595060



预训练网络

预训练网络是一个保存好的之前已在大型数据集（大规模图像分类任务）上训练好的卷积神经网络

预训练网络

如果这个原始数据集足够大且足够通用，那么预训练网络学到的特征的空间层次结构可以作为有效的提取视觉世界特征的模型。

预训练网络

即使新问题和新任务与原始任务完全不同

学习到的特征在不同问题之间是可移植的，这也是深度学习与浅层学习方法的一个重要优势。它使得深度学习对于小数据问题非常的有效。

Pytorch内置预训练网络

Pytorch库中包含

VGG16、VGG19、densenet

ResNet、mobilenet

Inception v3、

Xception等经典的模型架构。

ImageNet

ImageNet是一个手动标注好类别的图片数据库(为了机器视觉研究), 目前已有22,000个类别。

ImageNet

当我们在深度学习和卷积神经网络的背景下听到

“ImageNet”一词时，我们可能会提到ImageNet视觉识别比赛，称为ILSVRC。

ImageNet

这个图片分类比赛是训练一个模型，能够将输入图片正确分类到1000个类别中的某个类别。训练集120万，验证集5万，测试集10万。

ImageNet

这1,000个图片类别是我们在日常生活中遇到的，例如狗，猫，各种家居物品，车辆类型等等。

ImageNet

在图像分类方面，ImageNet比赛准确率已经作为计算机视觉分类算法的基准。自2012年以来，卷积神经网络和深度学习技术主导了这一比赛的排行榜。

VGG

VGG全称是Visual Geometry Group,属于牛津大学科学工程系，其发布了一些列以VGG开头的卷积网络模型，可以应用在人脸识别、图像分类等方面，分别从VGG16 ~ VGG19。

VGG16与VGG19

在2014年，VGG模型架构由Simonyan和Zisserman提出，
在“极深的大规模图像识别卷积网络”（Very Deep
Convolutional Networks for Large Scale Image
Recognition）这篇论文中有介绍

VGG

VGG研究卷积网络深度的初衷是想搞清楚卷积网络深度是如何影响大规模图像分类与识别的精度和准确率的，最初是VGG-16, 号称非常深的卷积网络全称为(VGG-Very-Deep-16 CNN)。

VGG16与VGG19

VGG模型结构简单有效，前几层仅使用 3×3 卷积核来增加网络深度，通过max pooling（最大池化）依次减少每层的神经元数量，最后三层分别是2个有4096个神经元的全连接层和一个softmax层。

VGG16与VGG19

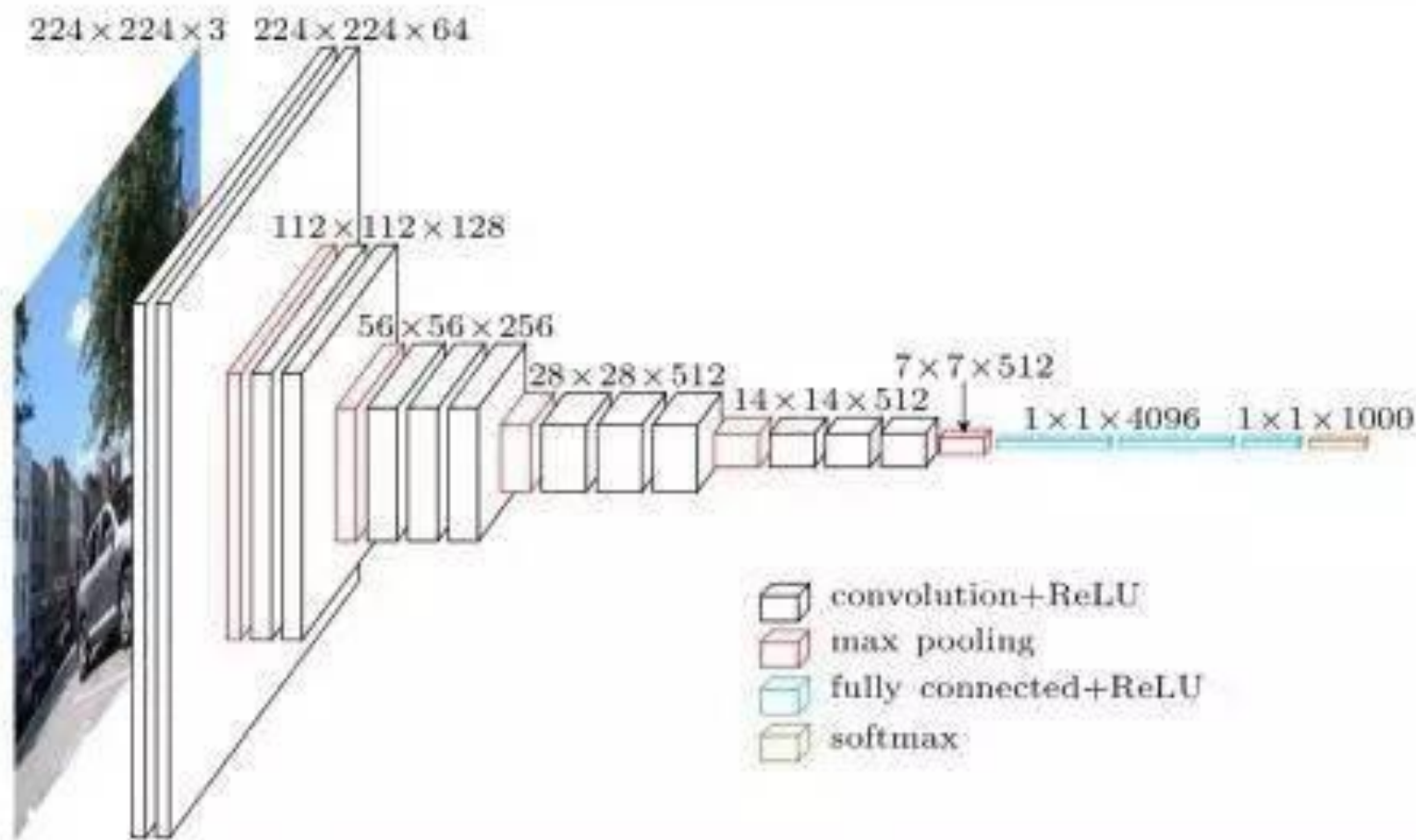


Figure 1: A visualization of the VGG architecture (source).

VGG

VGG在加深网络层数同时为了避免参数过多，在所有层都采用 3×3 的小卷积核，卷积层步长被设置为1。VGG的输入被设置为 224×224 大小的RGB图像，在训练集图像上对所有图像计算RGB均值，然后把图像作为输入传入VGG卷积网络，使用 3×3 或者 1×1 的filter，卷积步长被固定1。

VGG

VGG全连接层有3层，根据卷积层+全连接层总数目的不同可以从VGG11 ~ VGG19，最少的VGG11有8个卷积层与3个全连接层，最多的VGG19有16个卷积层+3个全连接层，此外VGG网络并不是在每个卷积层后面跟上一个池化层，还是总数5个池化层，分布在不同的卷积层之下。

VGG

ConvNet Configuration					
A	A-LRN	B	C	D	E
11 weight layers	11 weight layers	13 weight layers	16 weight layers	16 weight layers	19 weight layers
input (224×224 RGB image)					
conv3-64	conv3-64 LRN	conv3-64 conv3-64	conv3-64 conv3-64	conv3-64 conv3-64	conv3-64 conv3-64
maxpool					
conv3-128	conv3-128	conv3-128 conv3-128	conv3-128 conv3-128	conv3-128 conv3-128	conv3-128 conv3-128
maxpool					
conv3-256 conv3-256	conv3-256 conv3-256	conv3-256 conv3-256	conv3-256 conv3-256 conv1-256	conv3-256 conv3-256 conv3-256	conv3-256 conv3-256 conv3-256 conv3-256
maxpool					
conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 conv1-512	conv3-512 conv3-512 conv3-512	conv3-512 conv3-512 conv3-512 conv3-512
maxpool					
conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 conv1-512	conv3-512 conv3-512 conv3-512	conv3-512 conv3-512 conv3-512 conv3-512
maxpool					
FC-4096					
FC-4096					
FC-1000					
soft-max					

VGG

conv表示卷积层

FC表示全连接层(dense)

Conv3 表示卷积层使用3x3 filters

conv3-64表示 深度64

maxpool表示最大池化

VGG

在实际处理中还可以对第一个全连接层改为 7×7 的卷积网络，后面两个全连接层改为 1×1 的卷积网络，这个整个VGG就变成一个全卷积网络FCN。

VGG有两个很大的缺点

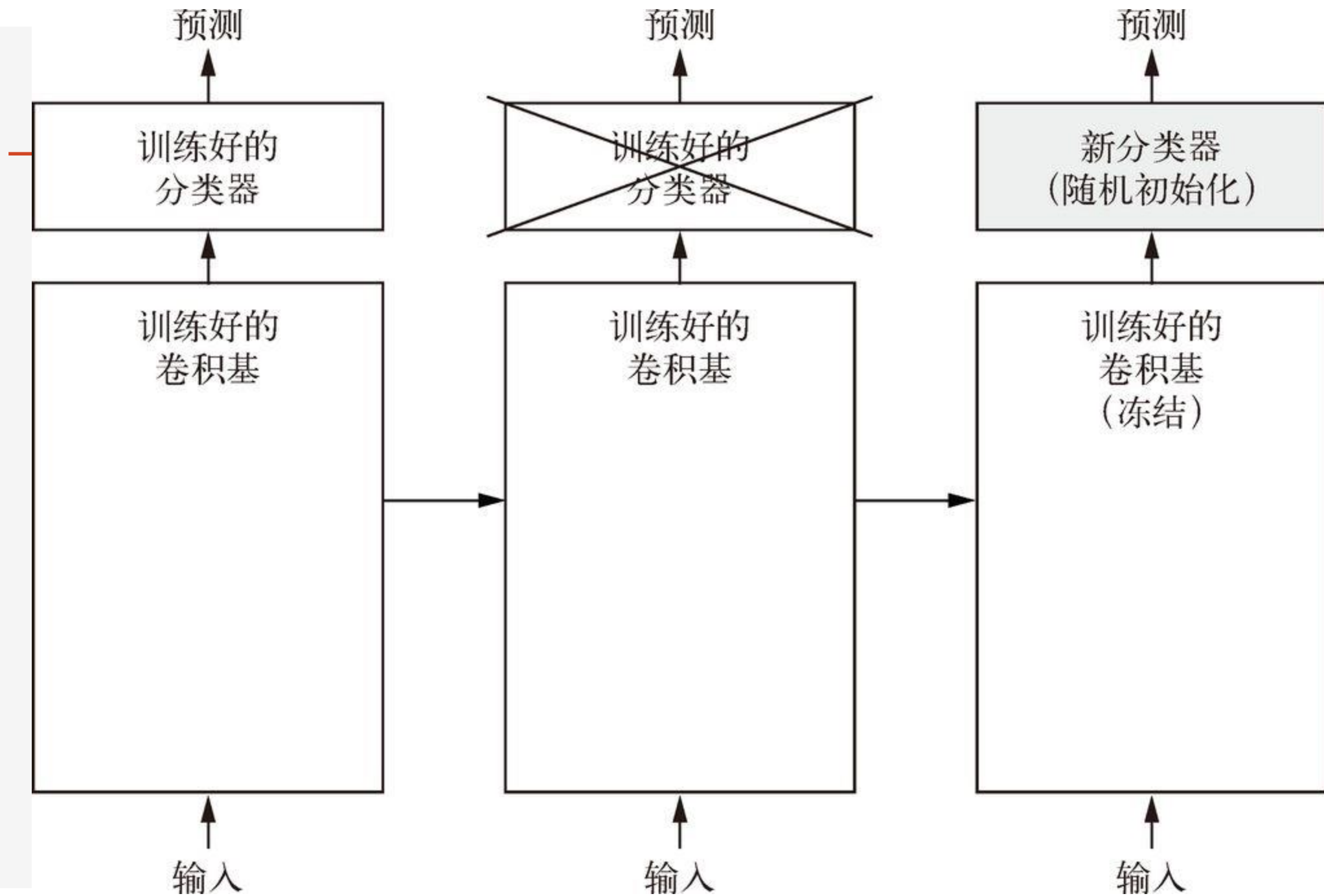
1. 网络架构weight数量相当大，很消耗磁盘空间。

2. 训练非常慢

由于其全连接节点的数量较多，再加上网络比较深，VGG16有533MB+，VGG19有574MB。这使得部署VGG比较耗时。

VGG

VGG在加深CNN网络深度方面首先做出了贡献，但是VGG也有自身的局限性，不能无限制的加深网络，在网络加深到一定层数之后就会出现训练效果退化、梯度消逝或者梯度爆炸等问题，总的来说VGG在刚提出的时候也是风靡一时，在ImageNet竞赛数据集上都取得了不错的效果



微调

所谓微调：

共同训练新添加的分类器层和部分或者全部卷积层。

这允许我们“微调”基础模型中的高阶特征表示，以使它们与特定任务更相关。。

微调

只有分类器已经训练好了，才能微调卷积基的卷积层。

如果有没有这样的话，刚开始的训练误差很大，微调之前这些卷积层学到的表示会被破坏掉

微调步骤

- 一、在预训练卷积基上添加自定义层
- 二、冻结卷积基所有层
- 三、训练添加的分类层
- 四、解冻卷积基的一部分层

五、联合训练解冻的卷积层和添加的自定义层

谢谢大家

讲师：日月光华 讲师微信：louhh01

