



南開大學
Nankai University

计算机学院
深度学习期末报告

卷积神经网络

姓名：杨馨仪
学号：2011440
专业：计算机科学与技术

2024 年 6 月 24 日

目录

1	实验要求	2
2	原版 CNN	2
3	微型 ResNet	2
4	微型 DenseNet	4
5	带有 Depthwise conv 的 MobileNets	5

1 实验要求

- 掌握 PyTorch 框架基础算子 (Convolution, Pooling, Activation) layers
- 学会使用 PyTorch 搭建简单的卷积神经网络来训练 Cifar10 数据集
- 了解经典的卷积神经网络结构

2 原版 CNN

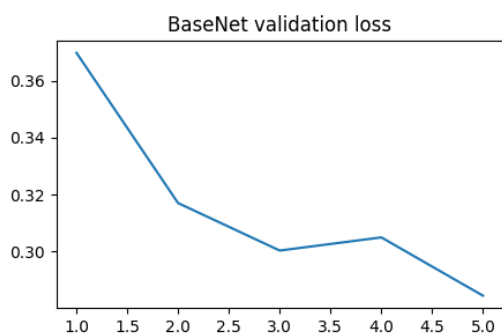
老师提供的原始版本卷积网络结构可用 `print(net)` 打印，结果如下：

```
Net(
  (conv1): Conv2d(3, 6, kernel_size=(5, 5), stride=(1, 1))
  (pool): MaxPool2d(kernel_size=2, stride=2, padding=0, dilation=1, ceil_mode=False)
  (conv2): Conv2d(6, 16, kernel_size=(5, 5), stride=(1, 1))
  (fc1): Linear(in_features=400, out_features=120, bias=True)
  (fc2): Linear(in_features=120, out_features=84, bias=True)
  (fc3): Linear(in_features=84, out_features=10, bias=True)
)
```

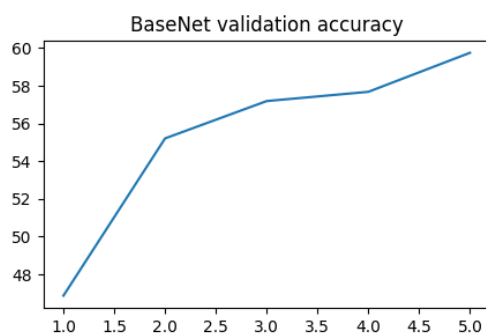
图 2.1: 原始版本卷积网络结构

这个 CNN 模型首先通过两个卷积层提取图像的空间特征，然后使用 ReLU 激活函数来引入非线性特性。每个卷积层之后紧跟一个 2×2 的最大池化层，其作用是进行下采样，减小特征图的尺寸，从而减少计算复杂度和内存占用。之后，网络将池化后的特征图展平，进入全连接层进行分类。最后一个全连接层输出 10 个节点，分别对应 CIFAR10 数据集的 10 个类别。

设置 `epochs=5`, `batchsize=4` 的情况下，在 Cifar 10 验证集上的训练 loss 曲线以及准确度曲线如下图所示：



(a) 原始版本 CNN 的损失曲线



(b) 原始版本 CNN 的准确度曲线

3 微型 ResNet

ResNet 的核心思想是通过引入残差块来构建网络。在每个残差块中，输入通过一个跳跃连接直接传递到输出，跳过了中间的卷积层和非线性层。这个跳跃连接允许梯度直接反向传播到前面的层，从而缓解了梯度消失的问题。

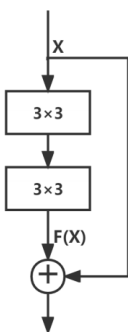


图 3.2: 一个基础残差块 (BasicBlock) 示意图

本实验中使用较简单的残差块 (BasicBlock) 实现了 Resnet34。其结果如下图所示:

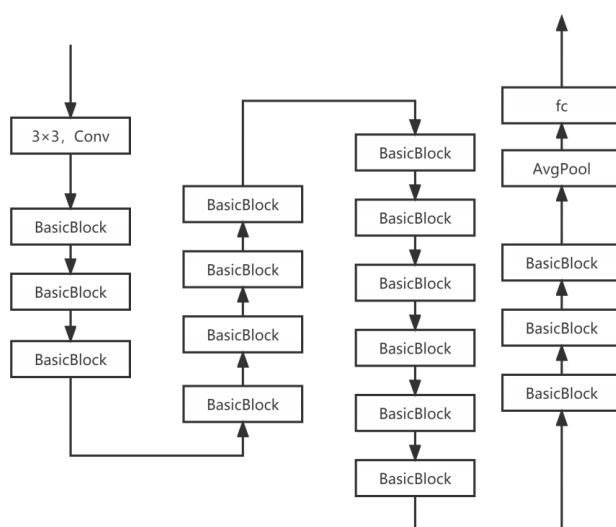
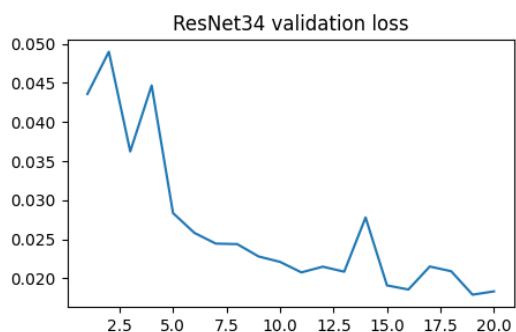
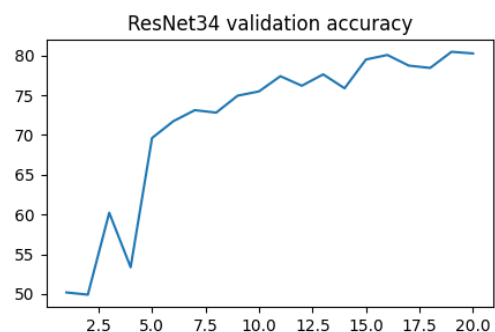


图 3.3: 个人实现的 ResNet34 网络结构

设置 epochs=20, batchsize=32, 使用 learning rate = 0.001 的 Adam 优化器的情况下, 在 Cifar 10 验证集上的训练 loss 曲线以及准确度曲线如下图所示:



(a) ResNet34 的损失曲线



(b) ResNet34 的准确度曲线

4 微型 DenseNet

DenseNet (Dense Convolutional Network) 是一种卷积神经网络架构，通过在各层之间建立密集连接，将所有层直接连接在一起来改进信息流动和梯度传递。其主要特征为：

- DenseNet 的基本构建模块是 Dense Block。在每个 Dense Block 中，所有层之间都存在直接连接。具体来说，第 l 层接收来自 0 到 $l-1$ 层的所有特征图。每个 Dense Block 的输出被传递到下一个 Dense Block 或过渡层。
- Dense Block 之间使用过渡层 (Transition Layer) 来进行降维和下采样，通常包含卷积层和池化层。过渡层的主要作用是减少特征图的数量和尺寸，从而控制模型复杂度和计算成本。

下图展示了 Dense Block 和 Transition Layer 的结构：

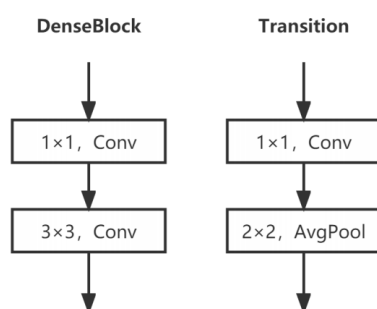


图 4.4: Dense Block 和 Transition Layer 示意图

进行模型训练时使用的是 DenseNet121 网络模型，其结构图如下：

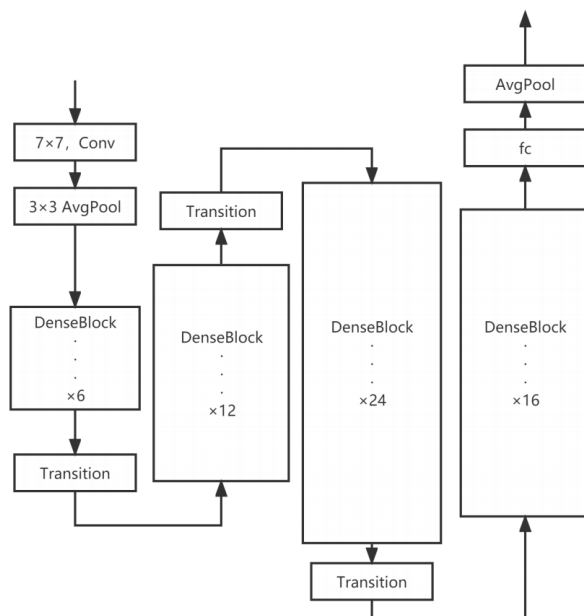
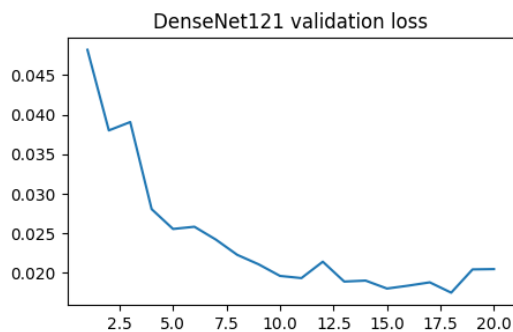
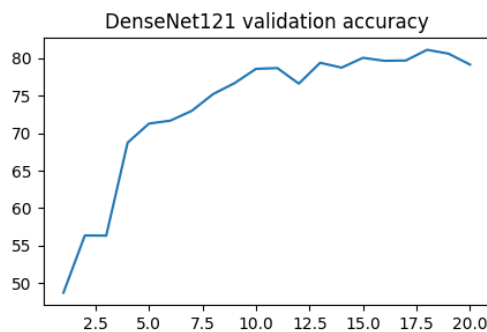


图 4.5: 个人实现的 DenseNet121 网络模型

设置 $\text{epochs}=20$, $\text{batchsize}=32$, 使用 $\text{learning rate} = 0.001$ 的 Adam 优化器的情况下，在 Cifar 10 验证集上的训练 loss 曲线以及准确度曲线如下图所示：



(a) DenseNet121 的损失曲线



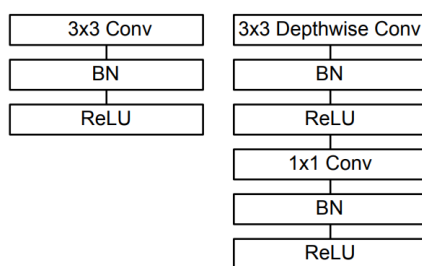
(b) DenseNet121 的准确度曲线

5 带有 Depthwise conv 的 MobileNets

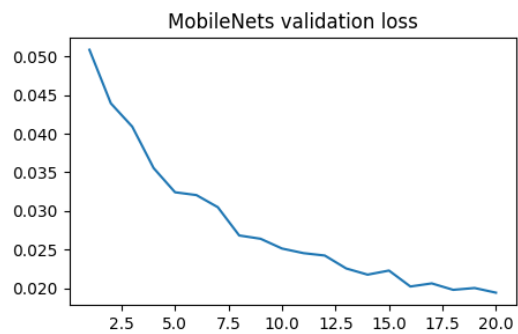
MobileNet 是一种轻量级卷积神经网络 (CNN) 架构, 专为移动和嵌入式设备的资源受限环境而设计。由 Google 在 2017 年提出, MobileNet 通过引入深度可分离卷积 (Depthwise Separable Convolutions) 来减少参数和计算量, 同时保持较高的准确性。其主要特点和优势如下:

- **深度可分离卷积:** 标准卷积分解为两个操作: 深度卷积和逐点卷积。深度卷积在每个输入通道上分别应用卷积核, 仅进行空间维度上的卷积。逐点卷积使用 1×1 卷积核在所有通道上进行卷积, 混合通道信息。这种分解显著减少了计算复杂度和参数量。
- **宽度乘数 (Width Multiplier, α):** 控制网络的宽度, 通过缩放每一层的通道数来调整模型的复杂度和计算需求。例如, $\alpha = 1$ 表示完整模型, $\alpha < 1$ 表示减少通道数的模型, 适合更加受限的环境。
- **分辨率乘数 (Resolution Multiplier, ρ):** 控制输入图像的分辨率, 通过缩放输入图像的尺寸来调整模型的计算需求。例如, $\rho = 1$ 表示原始分辨率, $\rho < 1$ 表示降低分辨率的模型。

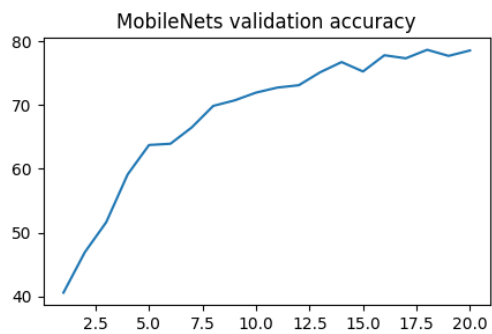
下图左侧为带有 batchnorm 和 ReLU 的标准卷积层, 右侧为深度可分离卷积 (深度层和点层) 其次是 batchnorm 和 ReLU 层。



设置 epochs=20, batchsize=32, 使用 learning rate = 0.001 的 Adam 优化器的情况下, 个人实现的带有 Depthwise conv 的 MobileNets 网络结构在 Cifar 10 验证集上的训练 loss 曲线以及准确度曲线如下图所示:



(c) MobileNets 的损失曲线



(d) MobileNets 的准确度曲线