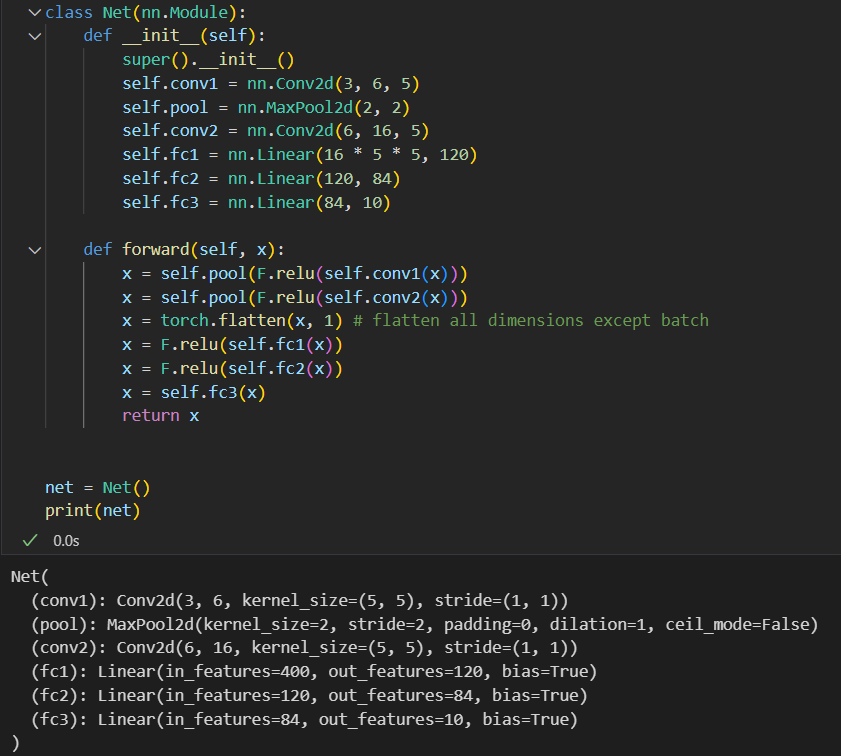
## 卷积神经网络实验报告

姓名：管昀玫 学号：2013750

### 实验要求

* 掌握卷积的基本原理
* 学会使用PyTorch搭建简单的CNN实现Cifar10数据集分类
* 学会使用PyTorch搭建简单的ResNet实现Cifar10数据集分类
* 学会使用PyTorch搭建简单的DenseNet实现Cifar10数据集分类
* 学会使用PyTorch搭建简单的SE-ResNet实现Cifar10数据集分类

### CNN



上图为老师给的CNN版本。注意，老师给的版本的forward函数有误，上图已进行了修正。该CNN结构解读如下：

1. 输入层是一个3通道的图像（RGB图像），因此输入的形状为(3, height, width)。
2. 第一个卷积层(conv1)有6个输出通道，使用5x5的卷积核进行卷积操作，卷积操作的步长为1。
3. 接下来是一个最大池化层(pool)，使用2x2的池化窗口进行池化操作，步长为2。这个池化层的作用是降低特征图的尺寸，提取主要特征并减少模型中的参数数量。
4. 第二个卷积层(conv2)有16个输出通道，同样使用5x5的卷积核进行卷积操作，步长为1。
5. 再次使用一个最大池化层(pool)，进行相同的2x2的池化操作，步长为2。
6. 紧接着是三个全连接层(fc1、fc2和fc3)，分别将特征图的展平后输入到这些全连接层中。fc1有120个输出神经元，fc2有84个输出神经元，fc3有10个输出神经元，对应于10个分类类别。

注意，老师给予的ipynb中有一句代码为：

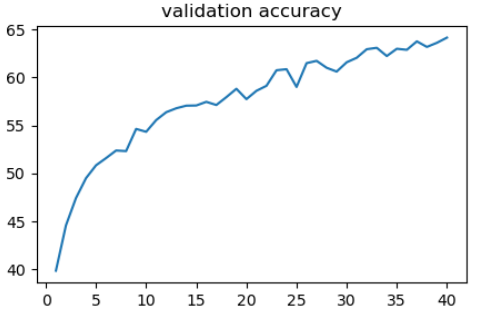
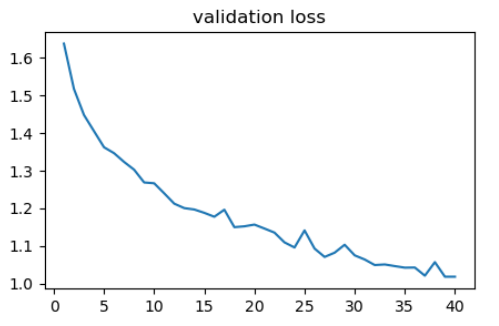
dataiter = iter(trainloader)

images, labels = dataiter.next()

由于pytorch版本不一致，在我本地应修改为：

images, labels = next(dataiter)

将该网络训练40轮，得到以下结果：



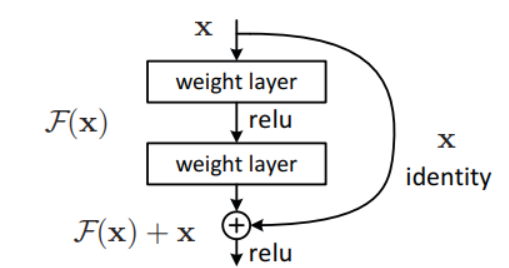
网络的准确率大约为64%左右。

### ResNet

随着网络深度的增加，精度变得饱和，然后迅速退化，这是因为网络可能收敛到局部最优，而非全局最优。resnet的提出主要是为了解决训练网络深度加深但准确率并没有提高，甚至不如浅层网络的问题。

为了解决第二个问题，论文作者提出了残差结构。在 ResNet 论文中，有两种映射方式：恒等映射（identity mapping）和残差映射（residual mapping）。恒等映射指的是通过右侧标有 "x" 的曲线来表示，而残差映射则指的是表示为 $F(x)$ 部分。最终的输出是 $F(x)$ 加上输入 $x$，即 $F(x) + x$。

ResNet 的优点在于，如果一个网络已经达到了最优状态，继续加深网络的层数可能会导致其他网络出现退化问题（错误率上升）。然而，对于 ResNet 网络来说，当网络已经接近最优状态时，残差映射中的 F(x) 部分会逐渐趋近于零，只剩下恒等映射部分。这意味着理论上网络将一直处于最优状态，网络的性能不会随着深度的增加而降低。这种残差映射的设计使得 ResNet 能够更好地训练非常深的神经网络，并获得更好的性能。



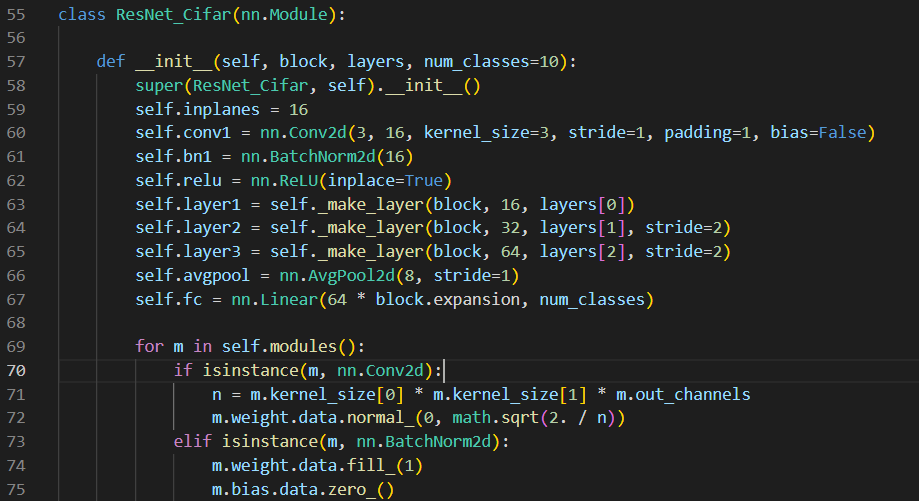
该残差块图源He K, Zhang X, Ren S, et al. Deep residual learning for image recognition[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2016: 770-778.

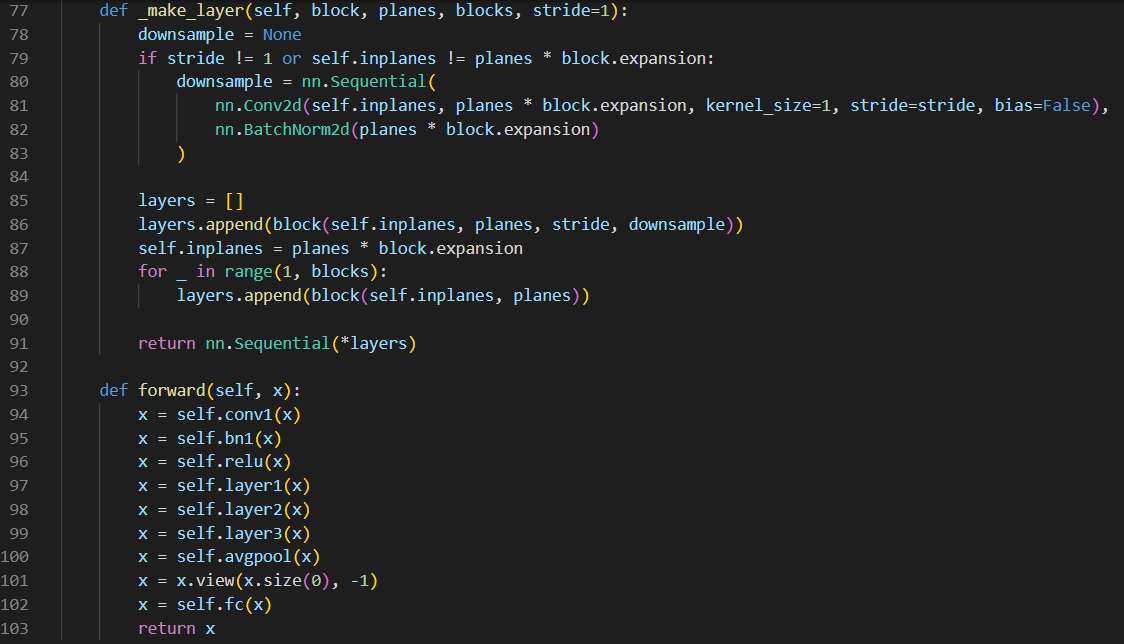
在论文中有两种残差块：

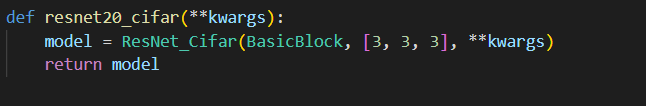
1. BasicBlock: 应用在较浅的网络中，如：Resnet 18和Resnet 34
2. BottleNeck: 应用在较深的网络中，如：Resnet 50 Resnet 101

我实现的resnet代码如下：

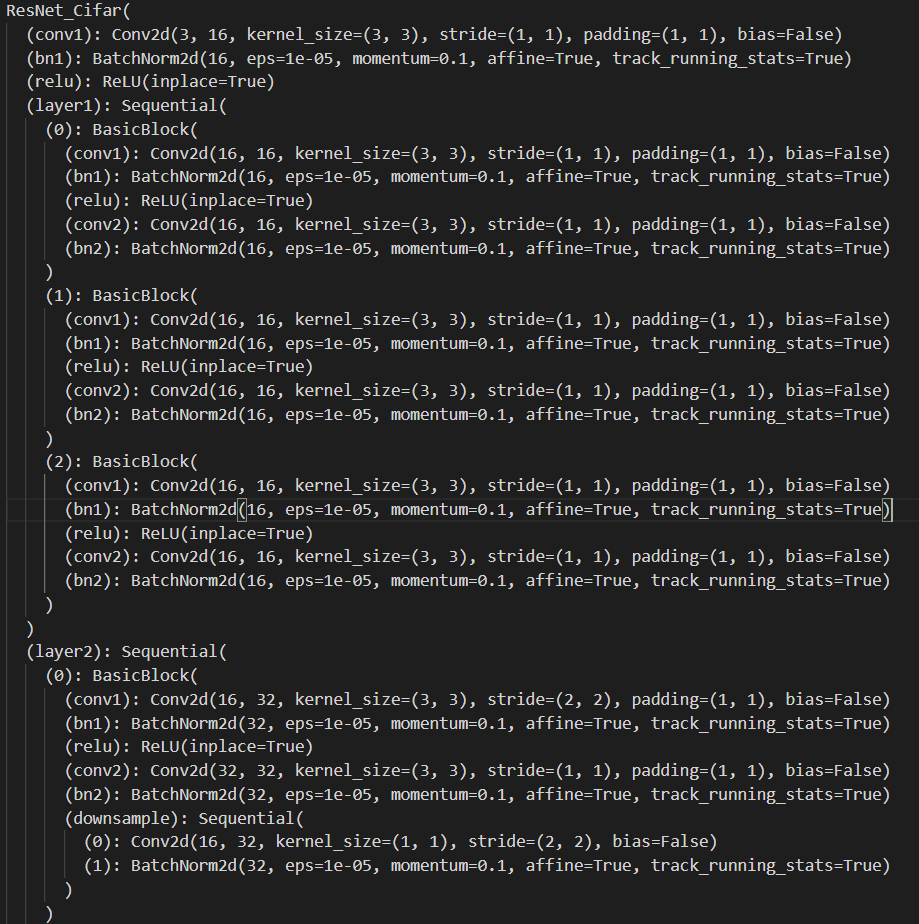






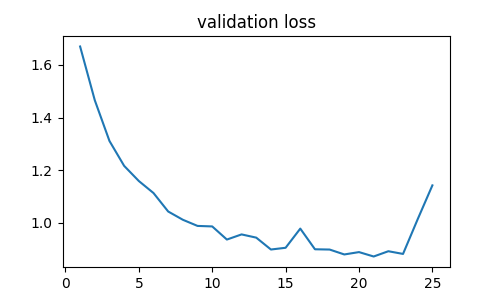
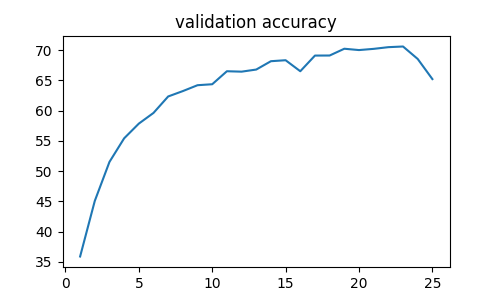


输入的数据x首先经过第一个卷积层、批归一化和ReLU激活函数，然后依次经过三个层（layer1、layer2、layer3），每个层包含3个残差块。接下来，通过平均池化层self.avgpool对特征进行空间降维，然后将其展平成一维向量，最后通过全连接层self.fc进行分类预测。



具体网络结构可见附件ResNet.txt文件。

将该网络训练50轮，得到以下结果：

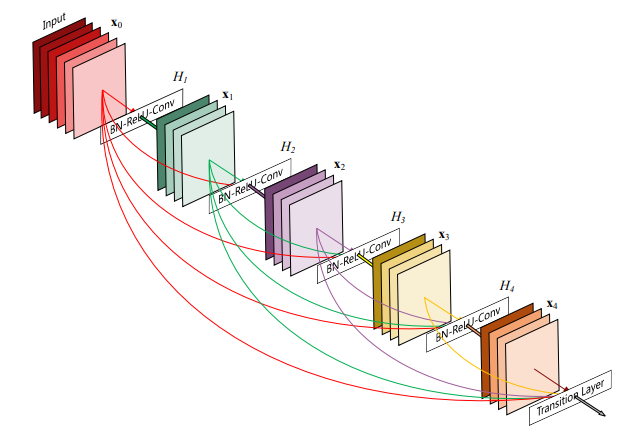


可以看见，在23轮左右，模型达到最佳，之后便开始过拟合，最高的准确率能达到70%左右，即使是20轮的准确率也超过了CNN，模型表现良好。

### DenseNet

与传统的卷积神经网络（CNN）不同，DenseNet通过密集连接的方式使得每一层都可以直接访问前面所有层的特征图。这种紧密连接的设计使得DenseNet在参数利用、特征重用和梯度传播等方面具有显著的优势。

DenseNet的核心思想是密集连接块（Dense Block）。在Dense Block中，每个层的输入来自前面所有层的拼接，这样每个层都可以直接接收到前面层的特征图作为输入。通过这种方式，网络中的信息传递更加充分，每个层都可以直接访问到前面层的特征，从而提高了特征的复用和网络的表达能力。



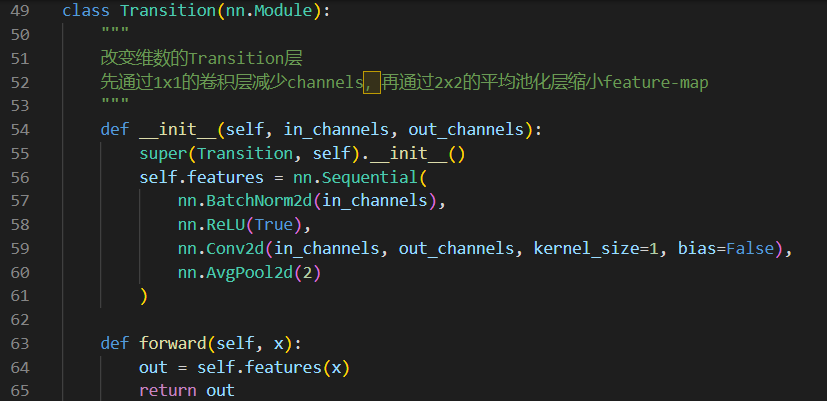
该图源自论文Huang G, Liu Z, Van Der Maaten L, et al. Densely connected convolutional networks[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2017: 4700-4708.

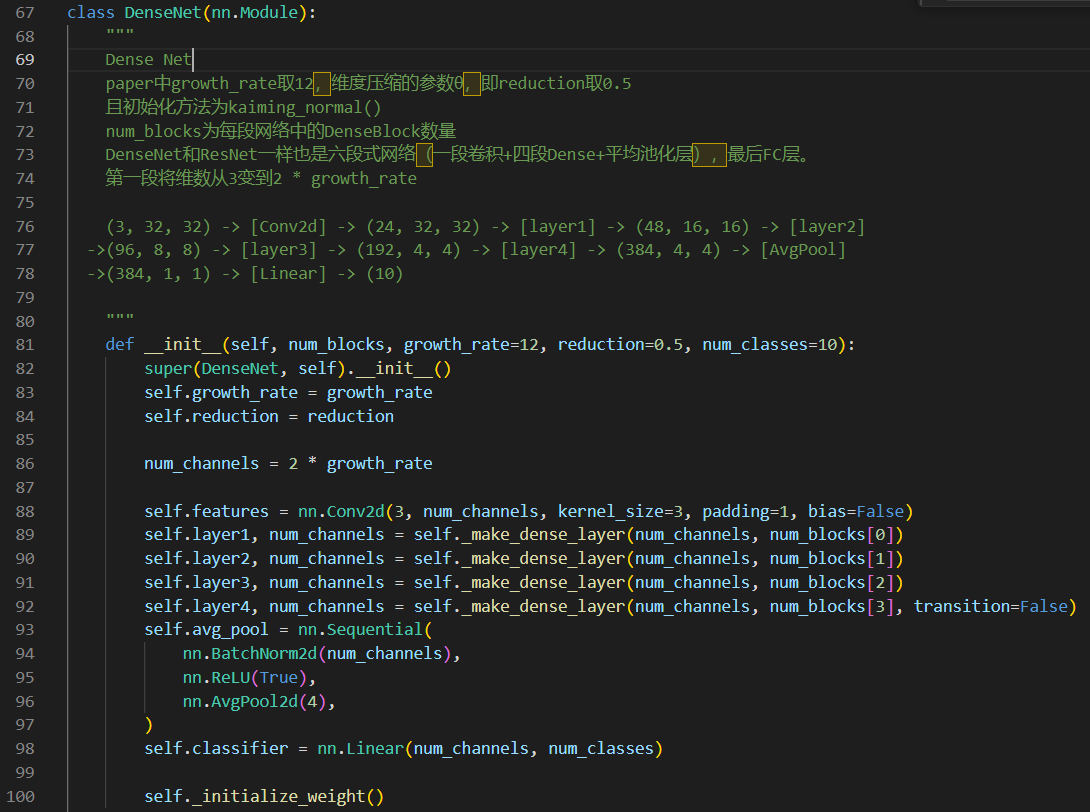
DenseNet有以下突出的特点：

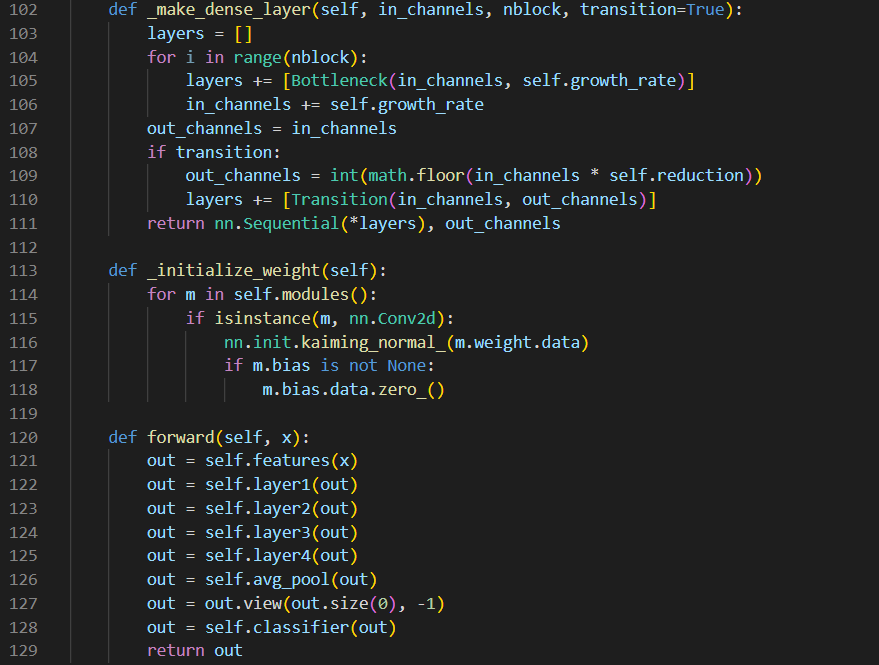
1. **密集连接块（Dense Block）**：Dense Block由多个卷积层组成，每个卷积层的输入是前面所有层的特征图的拼接。这种连接方式增加了网络的深度，同时使得每一层都能够接收到前面层的信息，有利于特征的传递和重用。
2. **过渡层（Transition Layer）**：在Dense Block之间，可以使用过渡层来减小特征图的维度。过渡层通常由一个卷积层和一个池化层组成，它可以减少特征图的尺寸和通道数，从而控制模型的复杂度并加速计算。
3. **稠密连接（Dense Connection）**：通过密集连接，每个层都可以直接接收到前面层的特征图，这种设计有助于特征的传播和信息的融合。相比于传统的网络结构，DenseNet减少了信息的丢失，并提高了特征的利用效率。
4. **全局特征复用**：由于每个层都可以直接访问前面层的特征，DenseNet具有较强的全局特征复用能力。这使得网络可以更好地捕捉到不同层次的特征，从而提升了模型的性能和泛化能力。

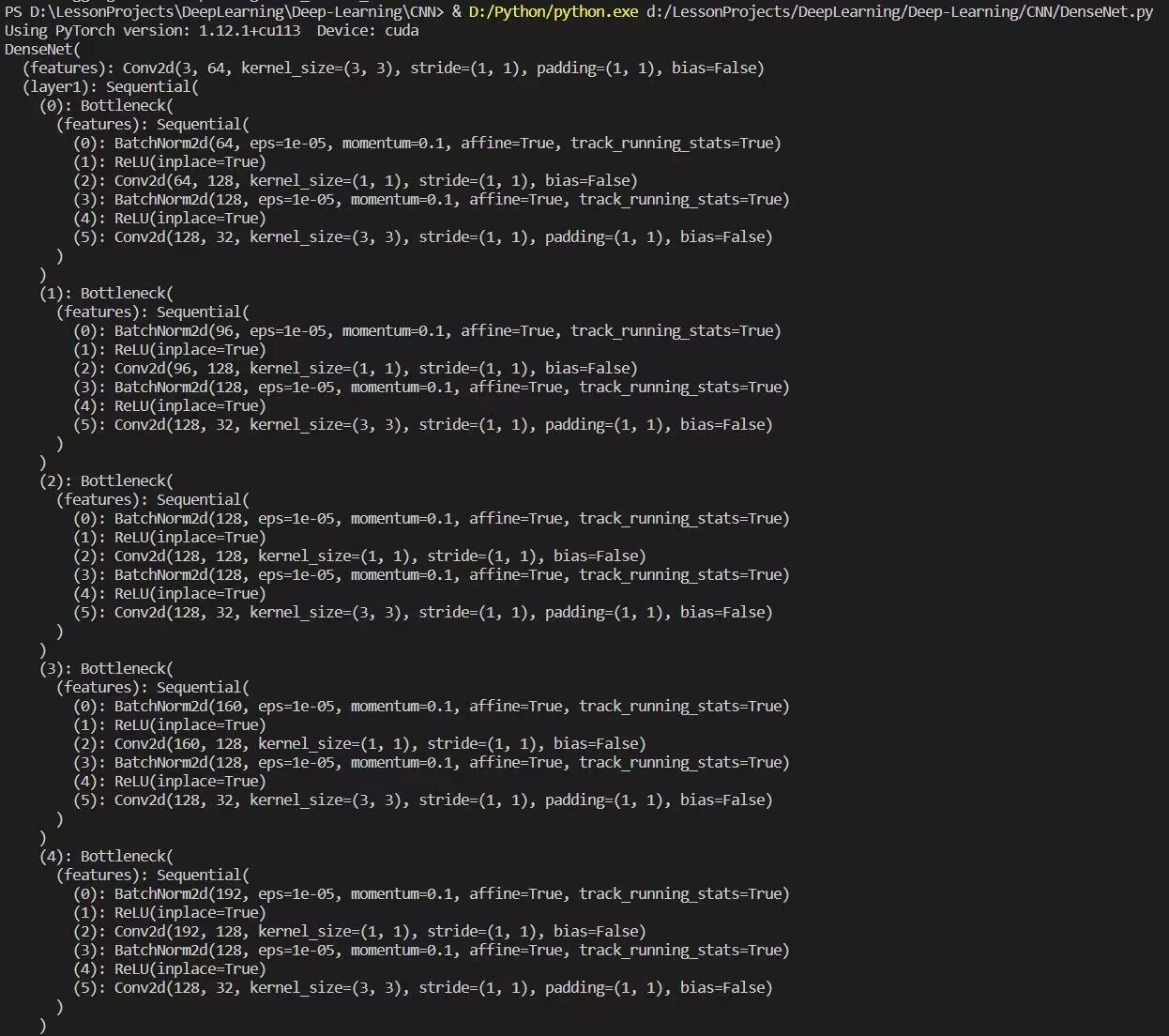
我实现的DenseNet如下所示：











具体网络结构可见附件DenseNet.txt文件。

该网络模仿DenseNet论文，growth\_rate取12，维度压缩的参数θ，即reduction取0.5

，且初始化方法为kaiming\_normal()。num\_blocks为每段网络中的DenseBlock数量。DenseNet和ResNet一样也是六段式网络（一段卷积+四段Dense+平均池化层），最后FC层。第一段将维数从3变到2 \* growth\_rate。

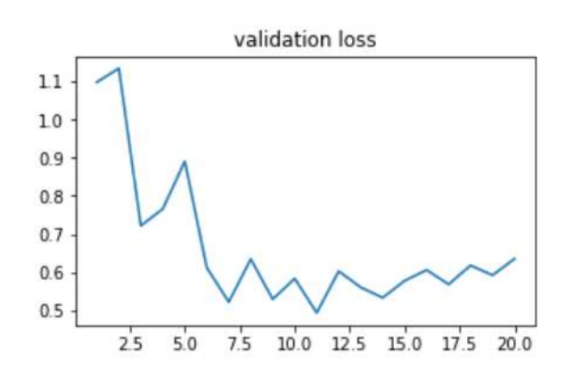
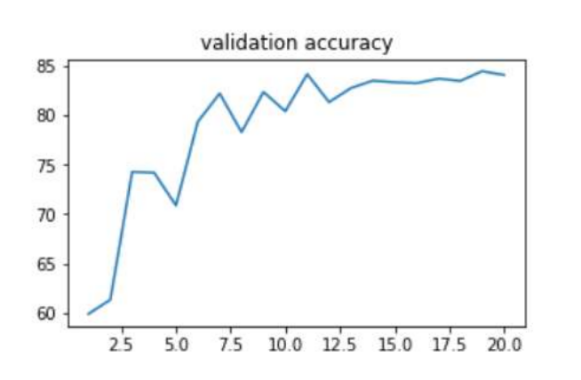
维度变化如下：

(3, 32, 32) -> [Conv2d] -> (24, 32, 32) -> [layer1] -> (48, 16, 16) -> [layer2]

->(96, 8, 8) -> [layer3] -> (192, 4, 4) -> [layer4] -> (384, 4, 4) -> [AvgPool]

->(384, 1, 1) -> [Linear] -> (10)

将该模型训练20epoch，得到以下结果：



20Epochs之后，模型已经收敛，正确率达到85%左右，高于ResNet。

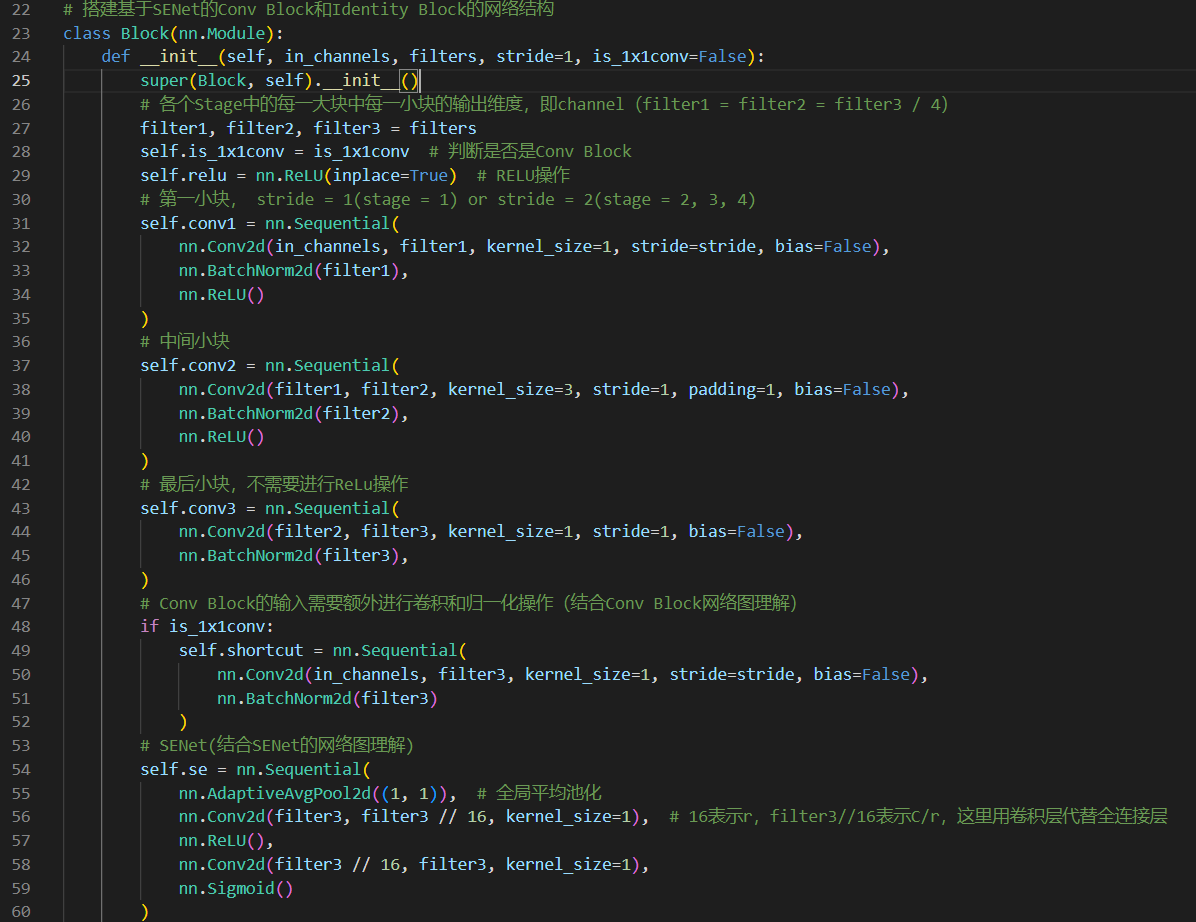
### SE-ResNet

卷积操作的设计包含spatial和channel-wise，过去的很多研究在spatial方面寻求更强大的表征力。SE模块则是考虑特征通道间的相关性。通过给不同的通道加权以强化更重要的特征，弱化不重要的特征通道，从而提高模型准确度。作者提出了Squeez and Excitation Blocks，这是一个通用的通道加权改造方法，很容易可以对过去的很多模型添加SE-Block结构。

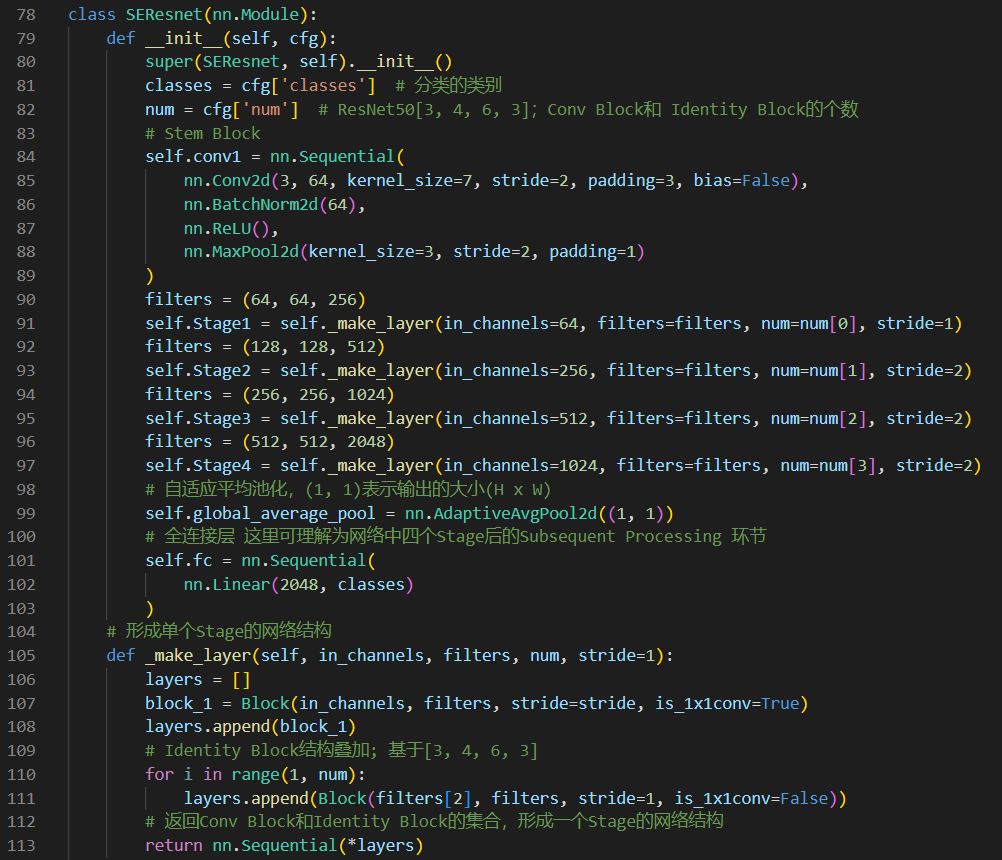
它的总体思路为：

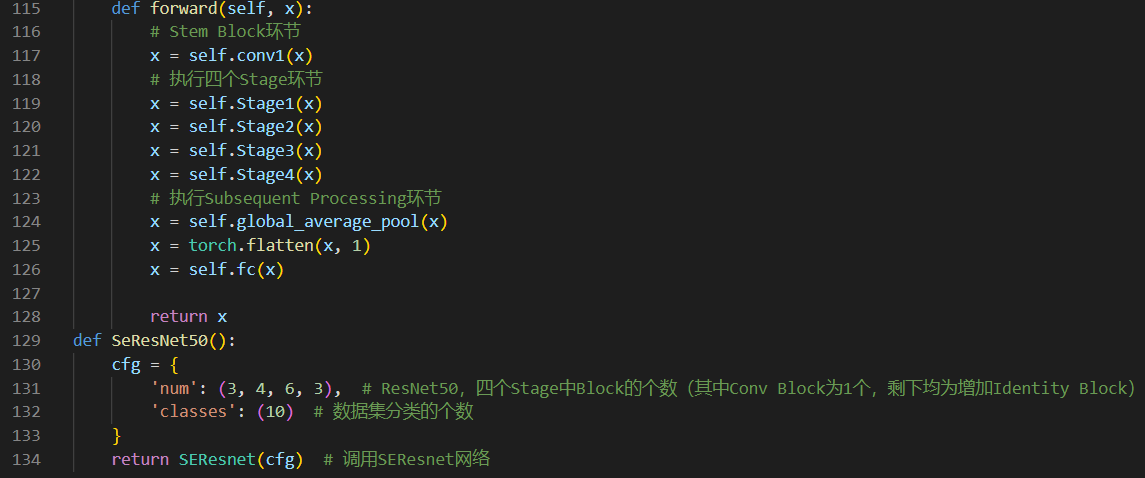
1. 引入了注意力的思想，即对于每个通道，用一个权重来表示该通道在下一阶段的重要性；
2. 做成了一个插入式的模块，十分方便与各个基础网络的结合。

我实现的SE-ResNet网络如下所示：

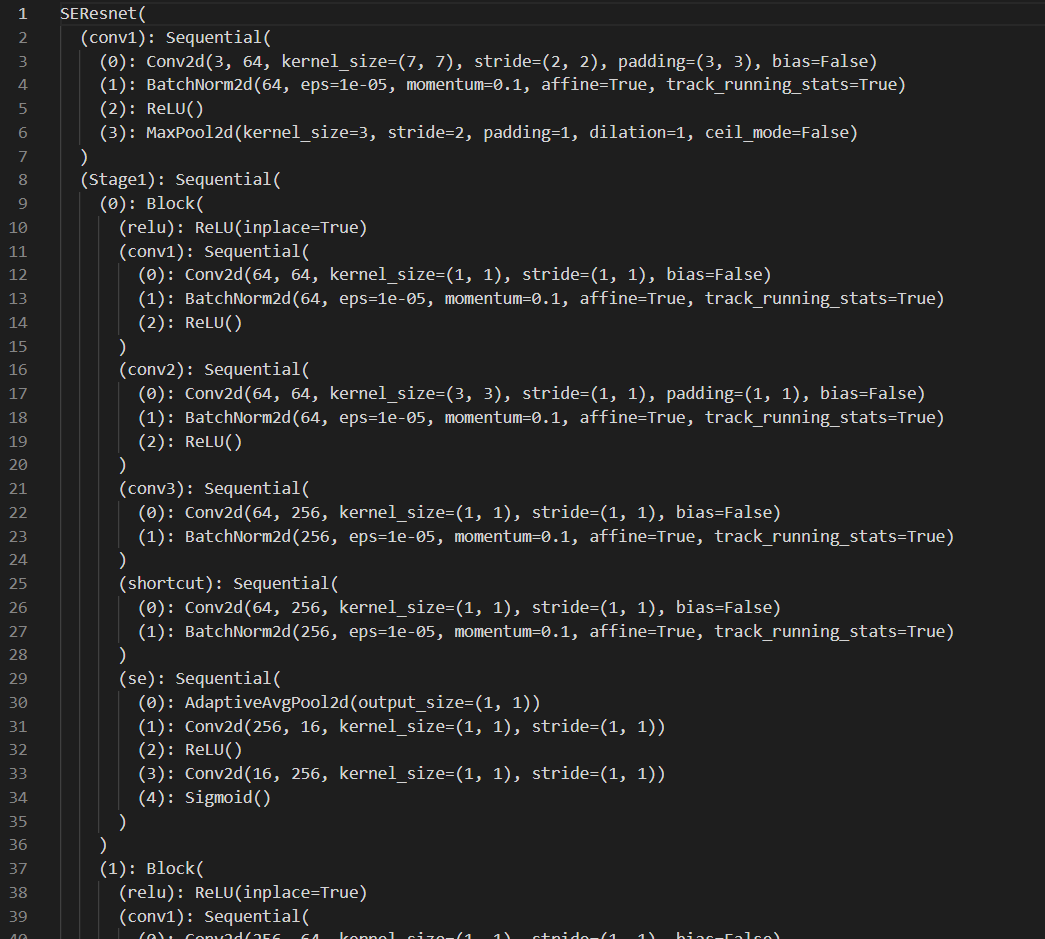






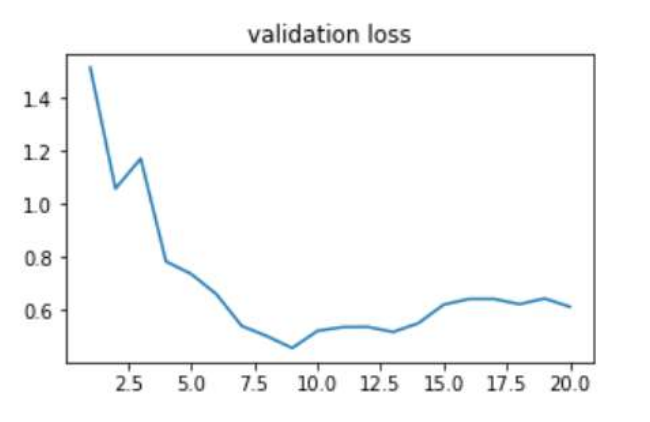
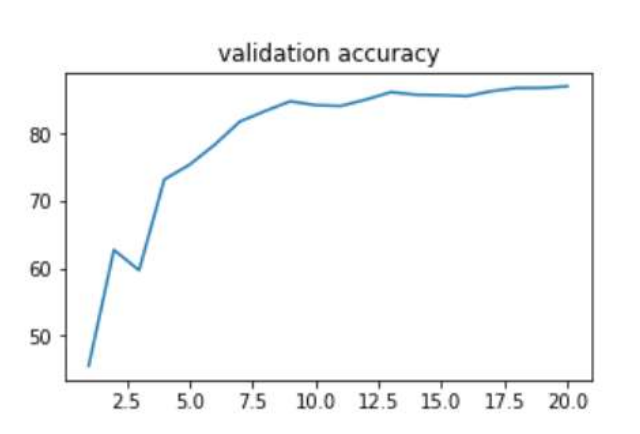


该网络的结构为：



具体网络结构可见附件SE-ResNet.txt。

将该网络训练20轮，得到以下结果：



其准确率大约为87%左右，为几个模型中的最佳。

### 没有跳跃连接的网络在训练过程中的不同

跳跃连接在卷积神经网络中的作用主要是解决梯度消失和梯度爆炸问题，并且有助于网络的训练和性能提升。梯度消失和梯度爆炸是由于梯度在反向传播时通过层数过多，与过多大于1或小于1的权重相乘所致。而如果在每个基本块中将输入与输出相加或连接，而这两个操作均不会在梯度前乘系数，故反向传播时梯度将通过跳跃连接直接越过该模块传导至上一模块，从而避免了梯度消失和梯度爆炸，使得神经网络的层数可以深达百层。从正向推理角度，具有跳跃连接的网络学习的不再是输入到输出的映射，而是输入到输出的残差。这样网络能够更轻松地学习到恒等映射，从而使得过参数模型也能取得良好效果。

1. 没有跳跃连接的卷积网络：

没有跳跃连接的卷积网络通过堆叠多个卷积层和池化层来提取图像特征。在训练过程中，使用反向传播算法来更新网络的参数，最小化损失函数。然而，由于梯度在反向传播过程中通过多个层级传递时容易消失或爆炸，这限制了网络的深度和性能。

1. ResNet：

ResNet引入了残差连接来解决梯度消失和梯度爆炸问题。残差连接是通过在网络中添加跨层的直接连接，将前一层的输出与后一层的输入相加。这样的设计允许信息在网络中直接跳跃传播，保留更多的梯度信息，有效地解决了梯度相关的问题。在训练过程中，ResNet通过最小化残差来进行参数更新，即最小化预测值与实际值之间的差异。

1. DenseNet：

DenseNet引入了稠密连接，通过将每个层的输出连接到后续所有层的输入上，实现了密集的信息流动。这样的连接方式使得网络中的每一层都能直接获得前面所有层的特征图作为输入，有效地促进了特征的传播和重用。在训练过程中，DenseNet通过最小化整个网络的损失函数来进行参数更新，同时所有层之间的连接也共享梯度信息，有助于更好地训练网络。

SE-ResNet：

SE-ResNet是在ResNet基础上引入了SE模块的改进版本。SE模块通过学习特征图中不同通道之间的关系来自适应地调整通道的重要性。它包括全局平均池化、一层全连接网络和sigmoid激活函数，通过学习通道的权重来增强有用的特征并抑制不相关的特征。在训练过程中，SE-ResNet通过最小化损失函数来同时更新ResNet中的参数和SE模块中的参数。

总体来说，ResNet通过残差连接实现跳跃连接，DenseNet通过稠密连接实现跳跃连接，而SE-ResNet在ResNet的基础上加入SE模块实现跳跃连接，并自适应地调整通道的重要性。这些改进方法在训练过程中的不同之处在于参数更新的方式和连接方式，从而使得网络能够更好地学习和优化特征。