1. **实验目的**

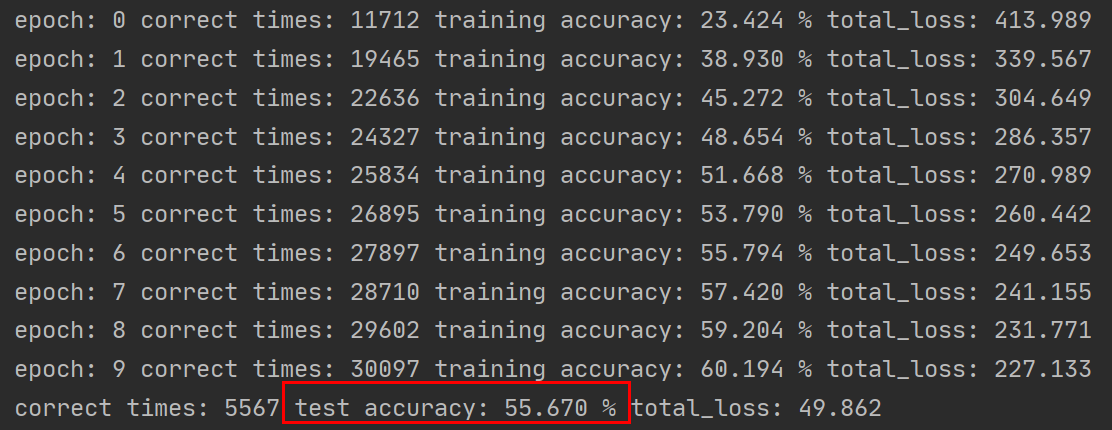
**(1) 熟悉新数据集 CIFAR-10，并和 MNIST 对比分类难度；**

**(2) 学习残差神经网络，特别是 Block 的概念；**

**(3) 构建残差神经网络，并基于此实现 CIFAR-10 的训练与测试。**

# 实验二

### **分析分类难度：CIFAR-10 vs MNIST**



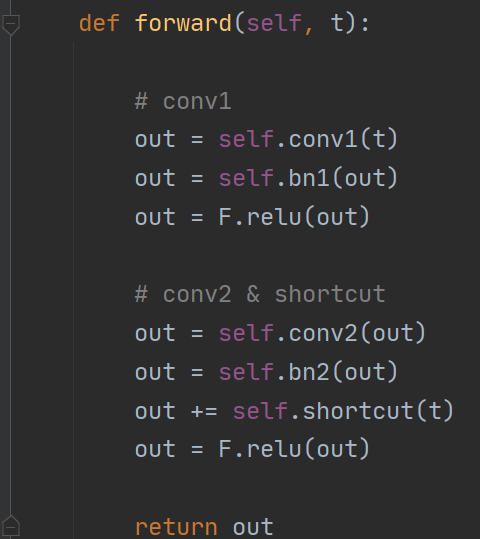
难度有了明显的提升。

#### 补全代码部分：

###### 一

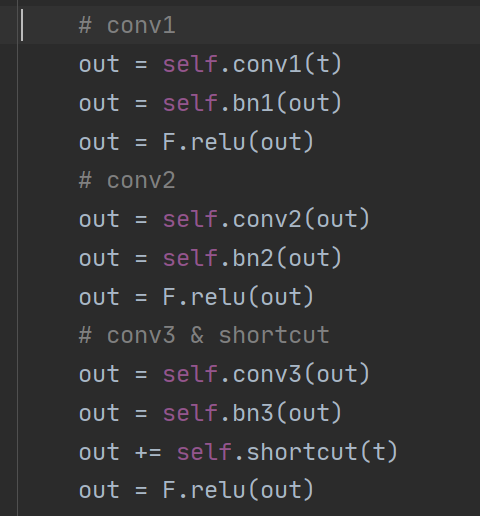
由残差网络结构可知这样补全：



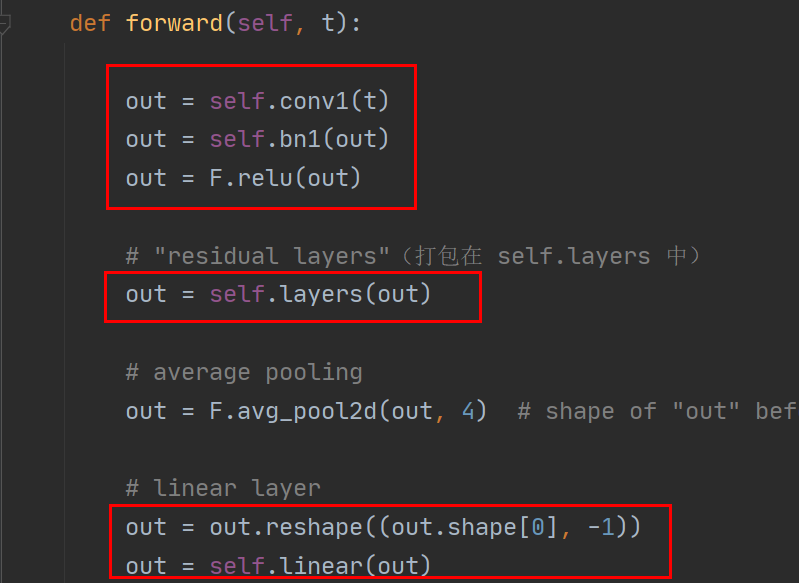


###### 二

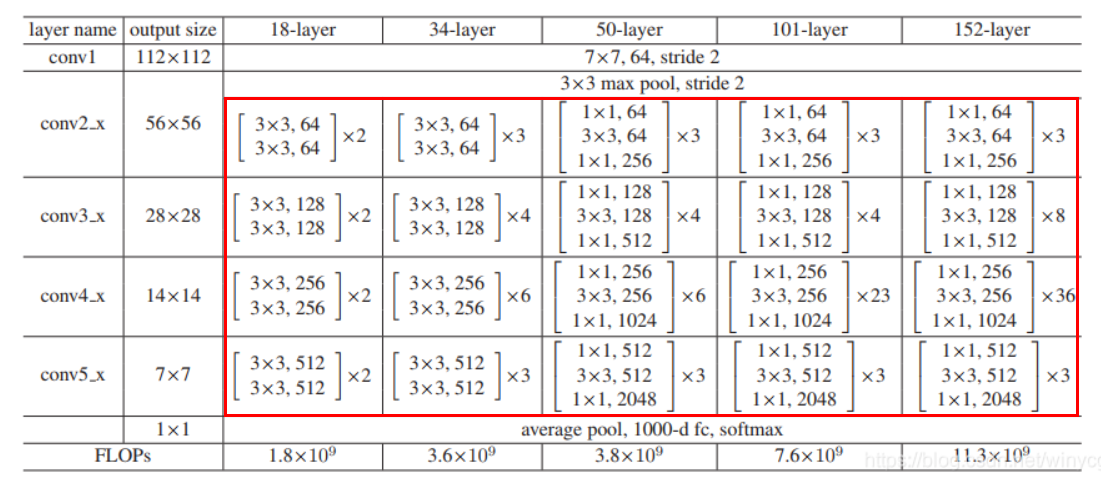


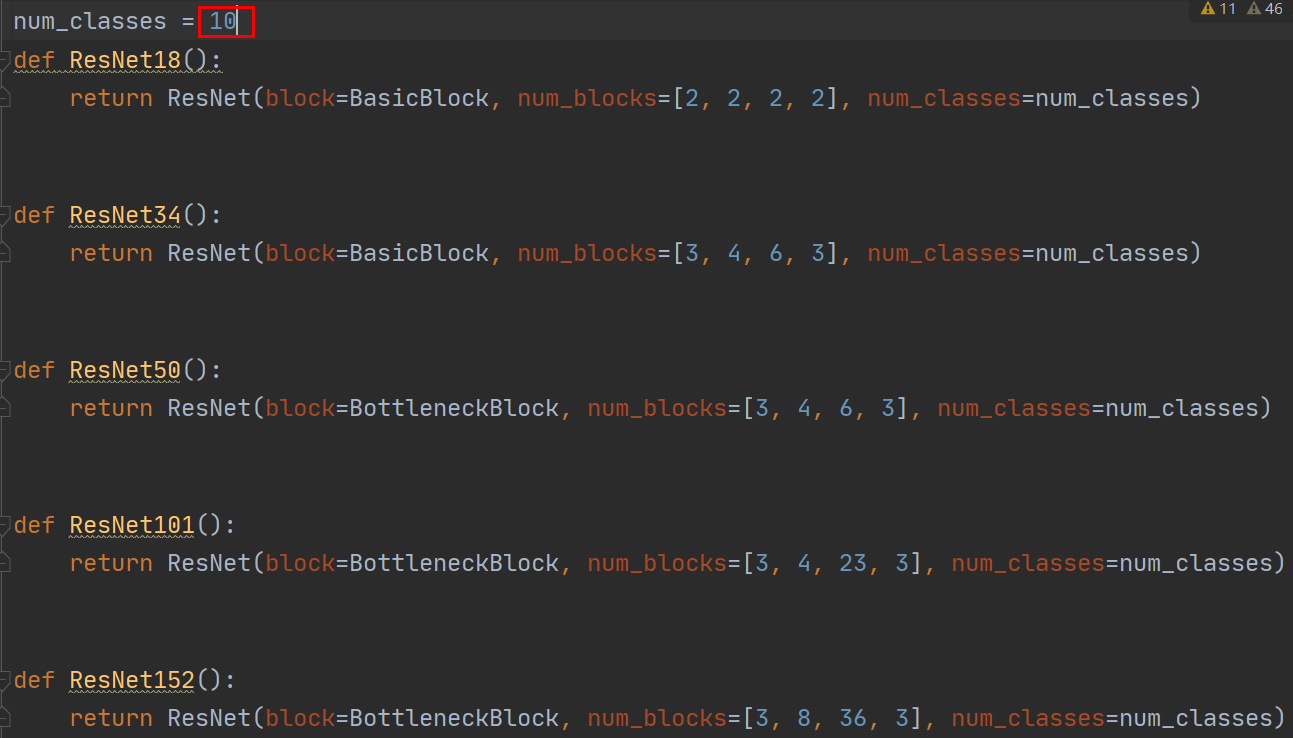


###### 三

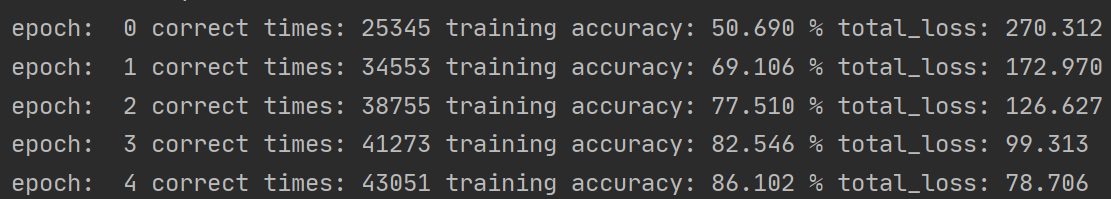


###### 四



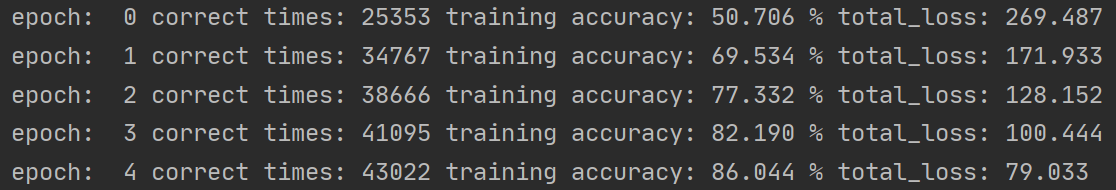


#### 展示结果：CIFAR-10 上训练并测试 ResNet18



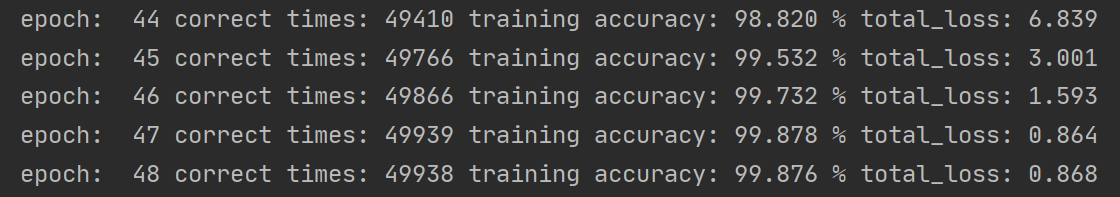


尝试：用resnet50训练五次

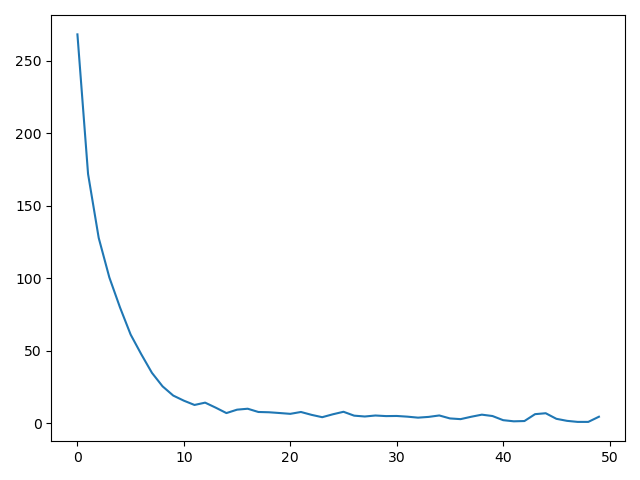




用resnet50训练，Epoch加大







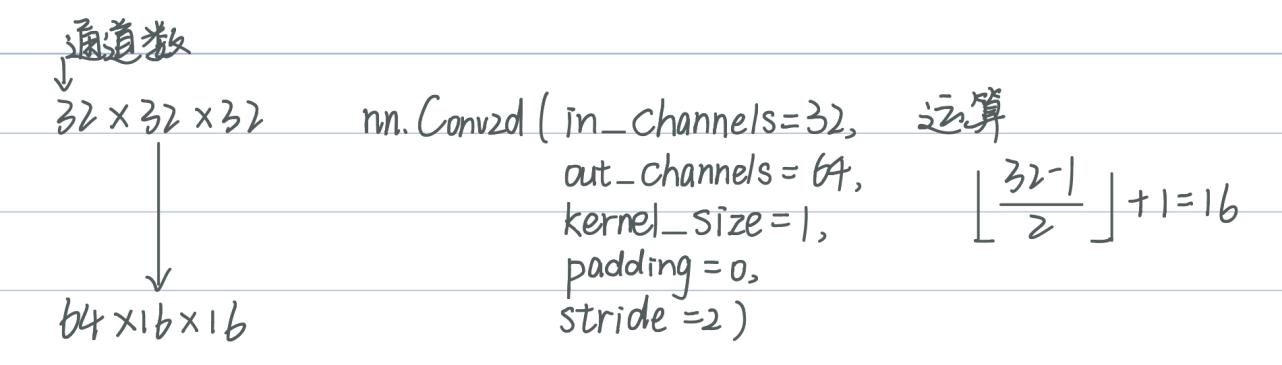
效果并没有很好的提升，说明不是epoch的层数，由图可知在epoch比较小的时候loss

已经比较小了，后期变化不大，需要用到网络优化的各种知识。（在后面补充的实验一中）

###### 请列出 ResNet18 第 3 个 shortcut connection 上的运算（在结构图中由虚线表示）及参与运算的张量形状，包括

###### 1) block 的输入在 shortcut connection 上经历 conv 变换前后的形状；

###### 2) 该 conv 变换的参数形状。

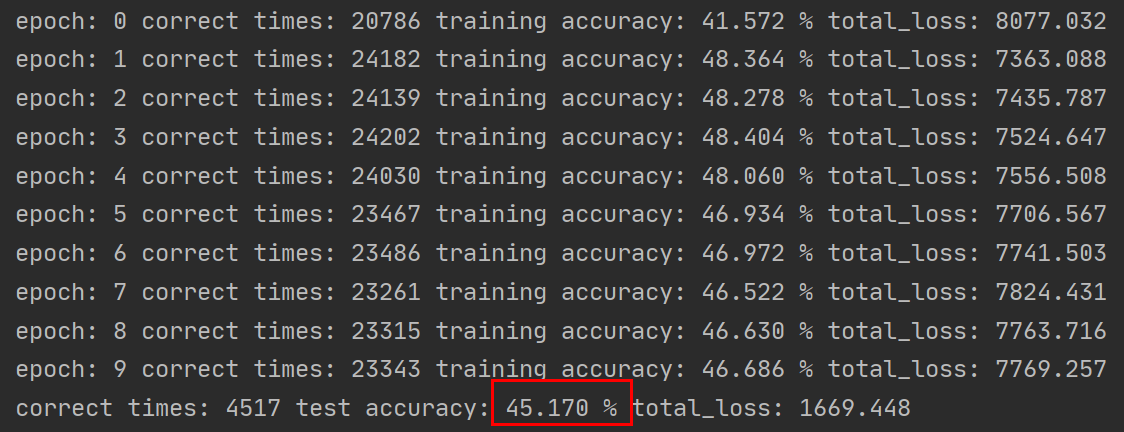


# 补充实验一的内容

****“4-1\_pytorch\_optimization.ipynb”****：

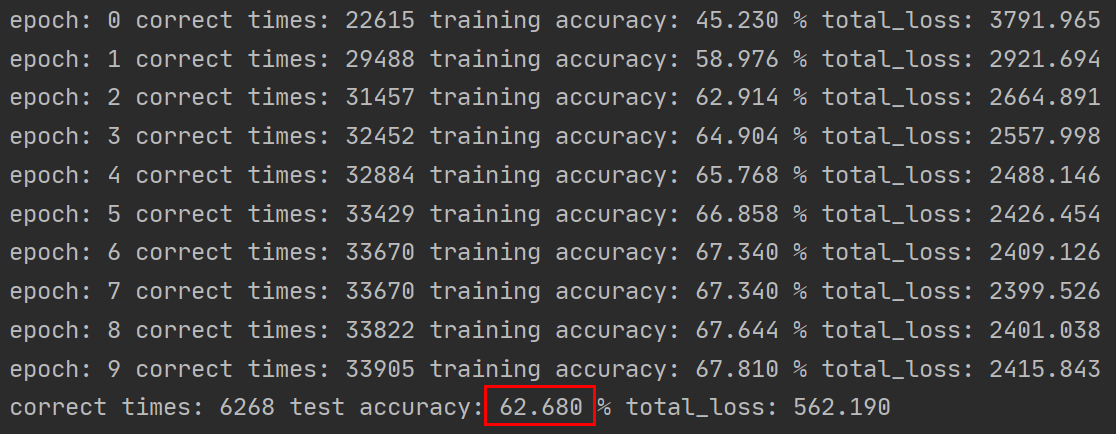
　　-　2. 对照组结果

　　　　请在实验报告中展示训练和测试结果，并讨论结果如何。如果再增加训练周期会对准确率有帮助吗？



对照组

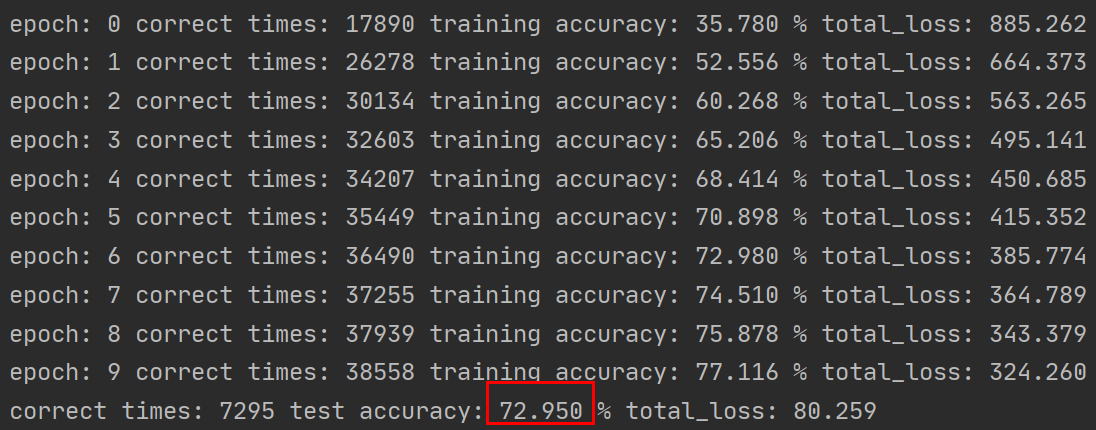
增加epoch到20，结果：



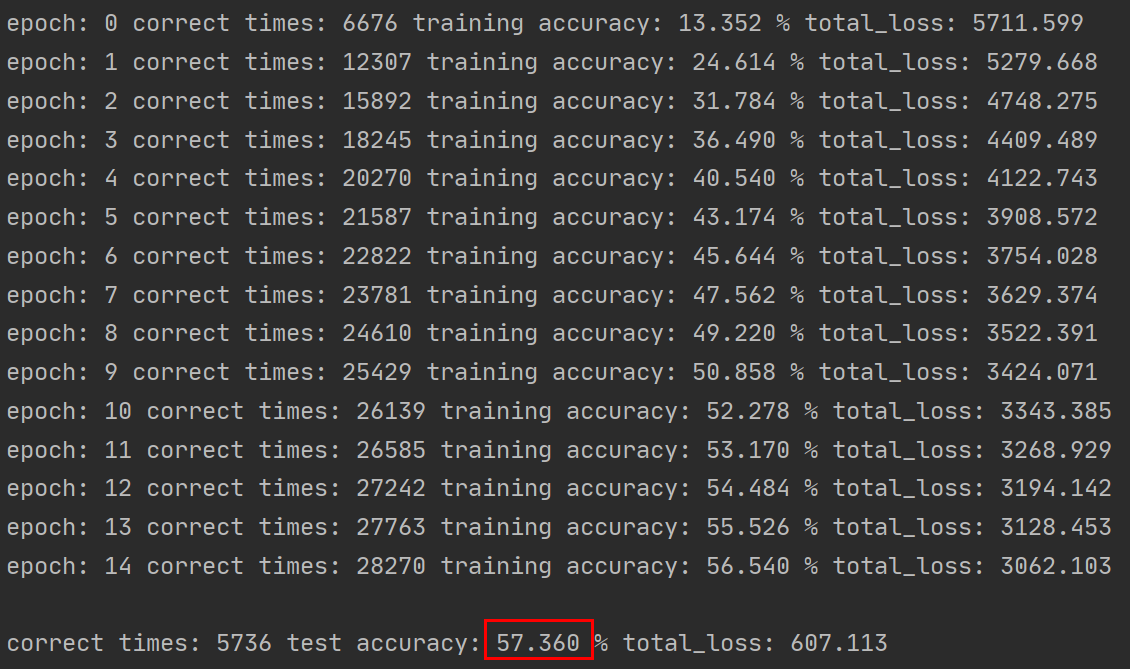
　　-　3. 小批量梯度下降

　　　　请同学们分别调整 batch size 和学习率，观察训练和测试结果并讨论。

将 batch size 从原先的 10 增加到 100，维持学习率 0.1 不变：



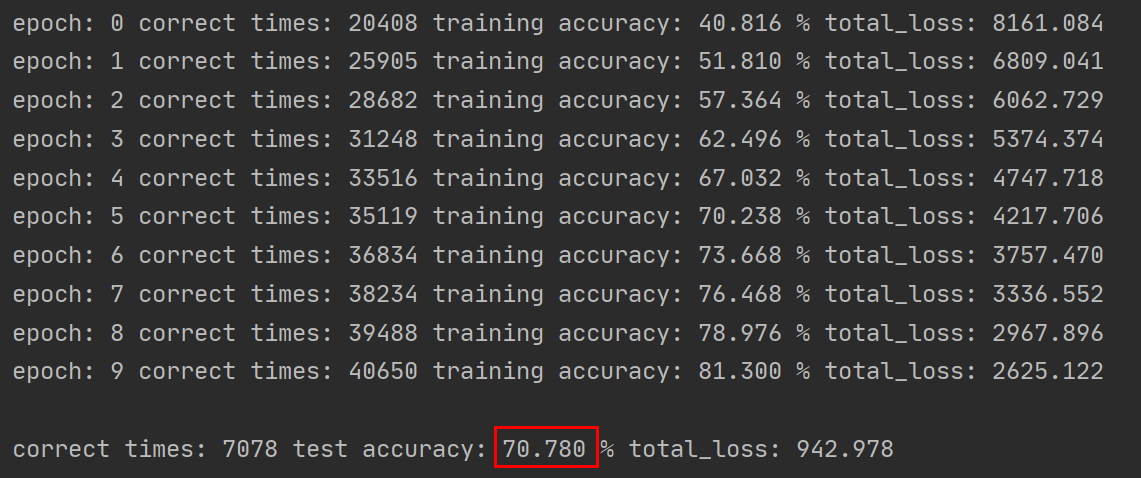
将 batch size 重新设置回 10，并将学习率从 0.1 减小至 0.001。由于学习率较小，训练较慢，可酌情增加训练周期数至 15：



结果：有所降低，结论：批量越大，随机梯度的方差越小，引入的噪声也越小，训练也越稳定，因此可以设置较大的学习率；批量较小时，需要设置较小的学习率，否则模型会不收敛。

　　-　4. 学习率衰减策略

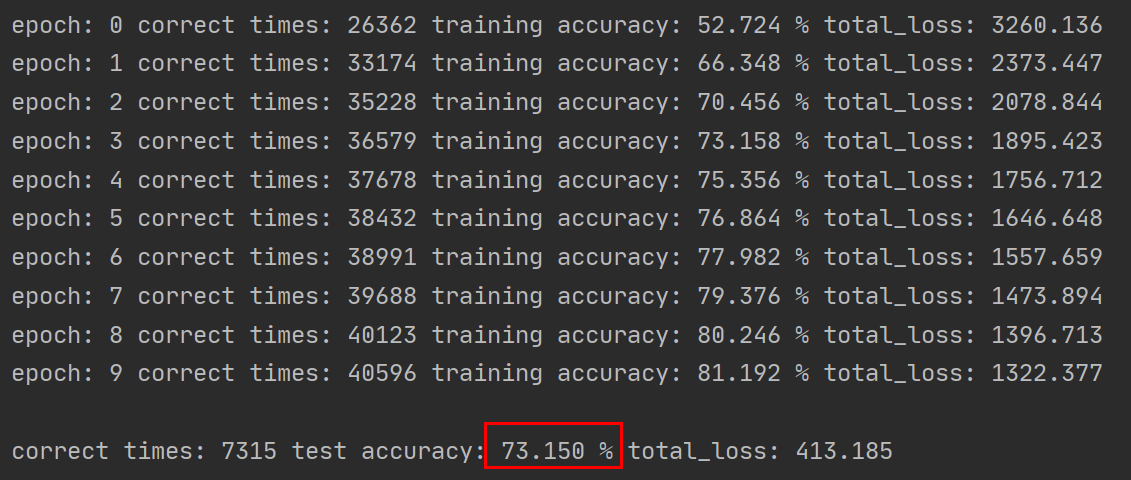
　　　　请同学们按照说明逐步减小学习率，观察训练和测试结果并讨论。



效果有所提升，学习率的选择对于最优解是非常有帮助的。学习率越大，输出误差对参数的影响就越大，参数更新的就越快，但同时受到异常数据的影响也就越大，很容易发散。学习率越小，模型收敛越慢，时间越长。理想的学习率是先大后小，存在衰减机制。

　　-　5. 批量归一化

　　　　请同学们实施批量归一化，观察训练和测试结果并讨论。

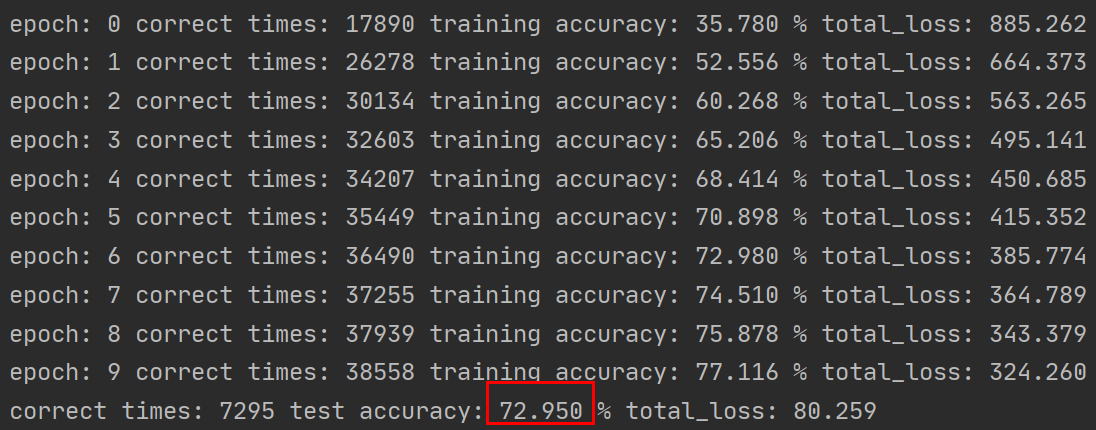


效果有所提升不加批量归一化的网络需要慢慢的调整学习率时，网络中加入批量归一化时，可以采用初始化很大的学习率，然后学习率衰减速度也很大，因此这个算法收敛很快。BN可以大大提高模型训练速度，提高网络泛化性能。数据批量归一化后相当于只使用了S型激活函数的线性部分，可以缓解S型激活函数反向传播中的梯度消失的问题。

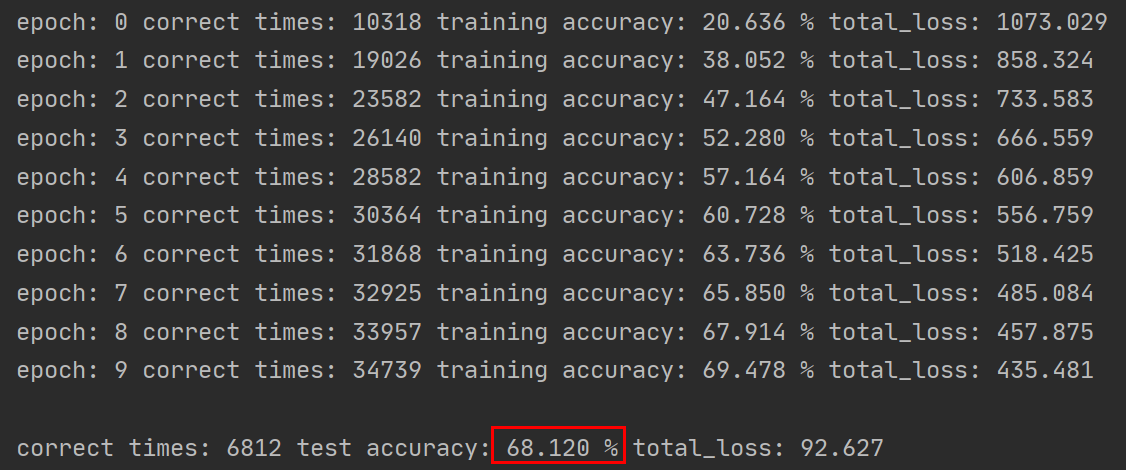
　　-　6. 数据预处理

　　　　请取消数据的标准化处理和 shuffle 设置，将结果和 3-(1) 对比，观察数据预处理是否对结果有影响。

3-(1)



消数据的标准化处理和 shuffle 设置的结果



有所下降，标准化是统一量纲，更方便训练的时候收敛，而shuffle是打乱数据，能够更好地实现无偏差训练模型，使得数据的分布更加均匀。

#### 实验心得

通过这次实验，我对残差神经网络的理解更加深刻了，并且也熟练了CIFAR-10这个数据集。在残差神经网络中，理解shortcut尤为关键，通过对网络的理解，再加上补全代码，实现残差网络的block，我更加深刻地理解参数一层一层的传导，进一步提升了代码能力，对pytorch网络搭建更加熟练了。并且我还学习了很多网络优化的方法，这个在训练模型上非常有用，以后需要优化网络的时候，懂得调整参数以及超参数，更有利于以后的发展。这是最后一次实验了，比较系统地过了整个神经网络与深度学习的内容，感觉收获颇多，对于一个问题，学会了怎么分析问题，搭建网络以及网络优化，希望以后能将学会的知识应用到实际中去，解决更多生活中亟待解决的问题。

