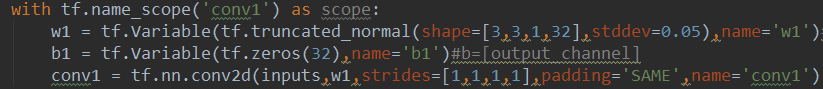
# 第十四周报告

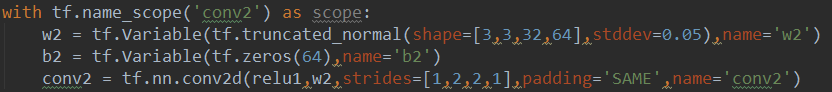
Siamese Network of MNIST

本周主要对于上一周的全连接模型进行改进。先进行3层卷积操作，之后进行全连接，同时在全连接层加入dropout和BN。

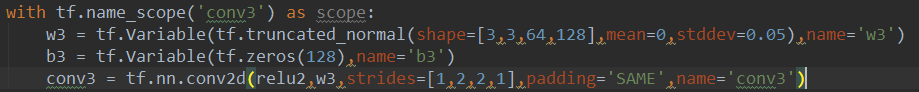
一、卷积层



以第一层卷积层为例，输入为（None，28，28，1）的张量，经过32个3\*3的卷积核，步长为1，padding为SAME，可以由公式得到输出长（宽）=输入长（宽）/步长，及输出为（None，28，28，32）的张量。



第二层，输入为第一层的输出，padding为SAME，步长为2，输出为（None，14，14，64）的张量。



第三层，输入为第二层的输出，padding为SAME，步长为2，输出为（None，7，7，128）的张量。

二、dropout（随机失活）

1、适用范围：tf.nn.dropout是TensorFlow里面为了防止或减轻过拟合而使用的函数，它一般用在全连接层。

2、原理：dropout就是在不同的训练过程中随机扔掉一部分神经元。也就是让某个神经元的激活值以一定的概率p，让其停止工作，这次训练过程中不更新权值，也不参加神经网络的计算。但是它的权重得保留下来（只是暂时不更新而已），因为下次样本输入时它可能又得工作了。



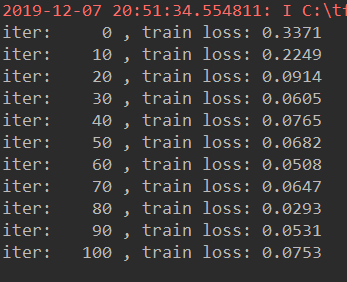
三、BN（批规范化操作）

2015年Google提出的Batch Normalization,不仅加快了模型的收敛速度，而且在一定程度上缓解了深层神经网络的“梯度弥散”问题，从而使得训练深层网络模型更加容易和稳定。另外，BN不仅适用于深层网络，对于传统的较浅层网络，批规范化操作也能对网络泛化性能起到一定的提升作用。目前批规范化操作已经成为几乎所有卷积神经网络的标配。



四、实验结果

经过100次学习率为0.01，batch\_size为64的训练之后得到的损失函数如下图。



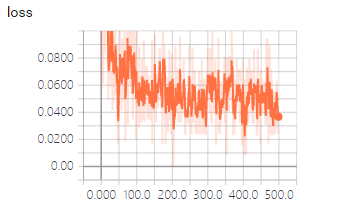


图1 改进之后损失函数 图2 改进之前损失函数

对比发现改进之后模型收敛更快，损失更小，但是时间比之前的模型要慢。  
总结：加入卷积层、BN层、dropout层能明显改善模型，但也会增加训练时间。

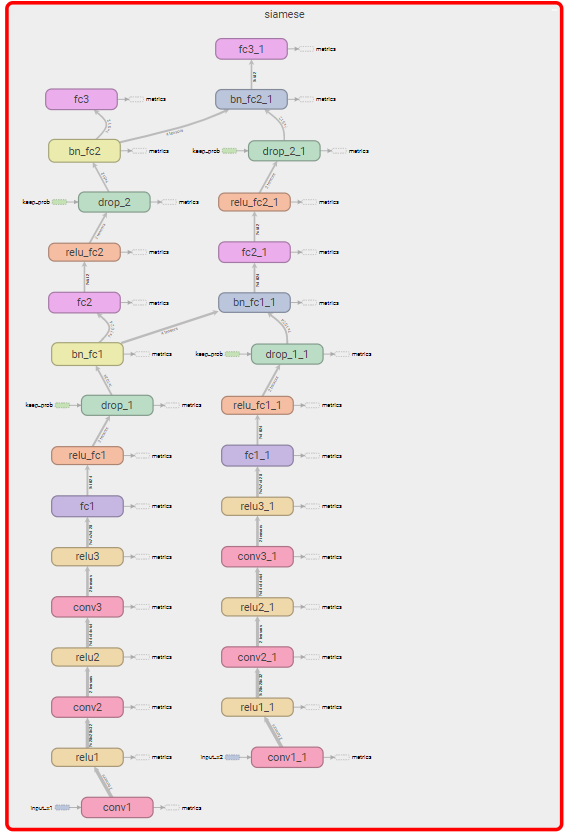


图3 改进之后的网络流图

报告人：宋政谦

时间：2019/12/08