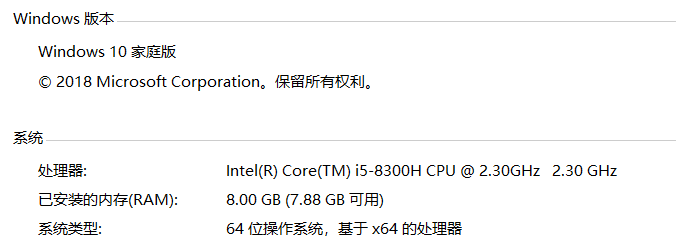
实验六 英文数字语音识别

**实验内容：**

基于TensorFlow框架及Python相关模块对语音信号进行特征提取，搭建声纹识别模型并训练，最后实现对英文数字语音zero-nine的识别。

**实验环境：**

Lenovo Legion Y7000p，Intel 8代i5，8G内存，显卡：NVIDIA GeForce GTX 1050 Ti 4G

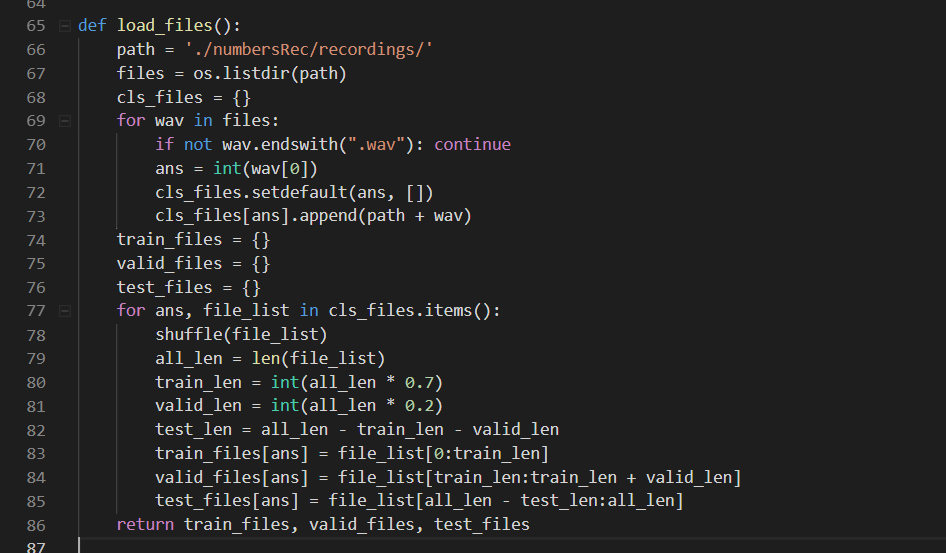


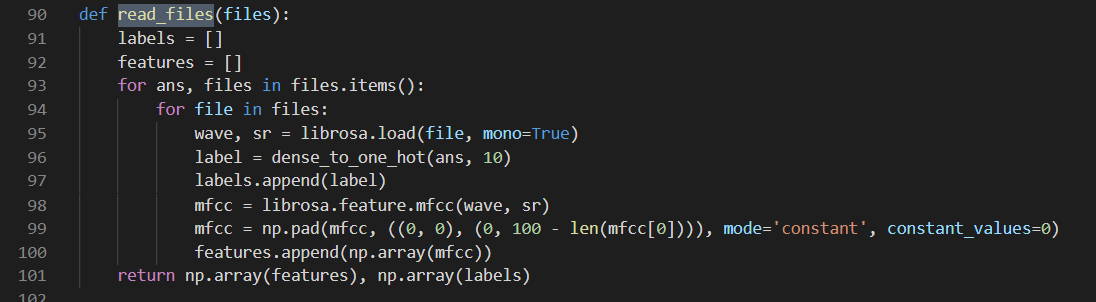
Python 3.6 + TensorFlow 1.11.0 + CUDA 9.0 + cuDNN 7.6.0 + VSCode

**实验步骤：**

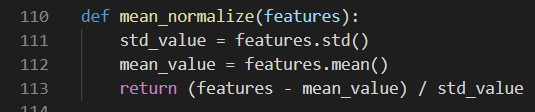
1. 数据预处理：

（1）加载文件并对对音频文件特征MFCC进行提取。

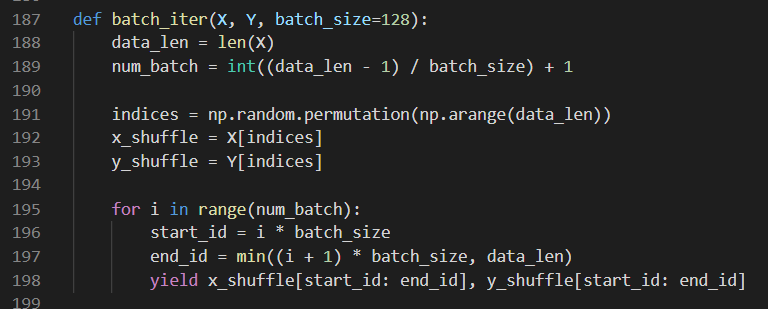




（2）对特征向量进行归一化处理

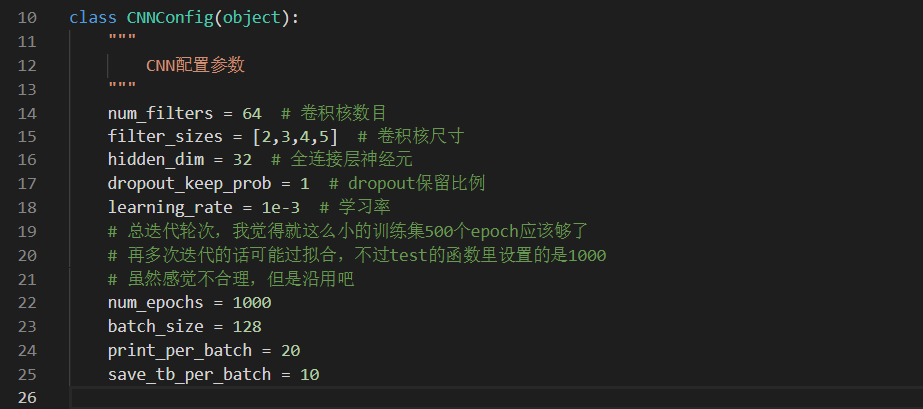


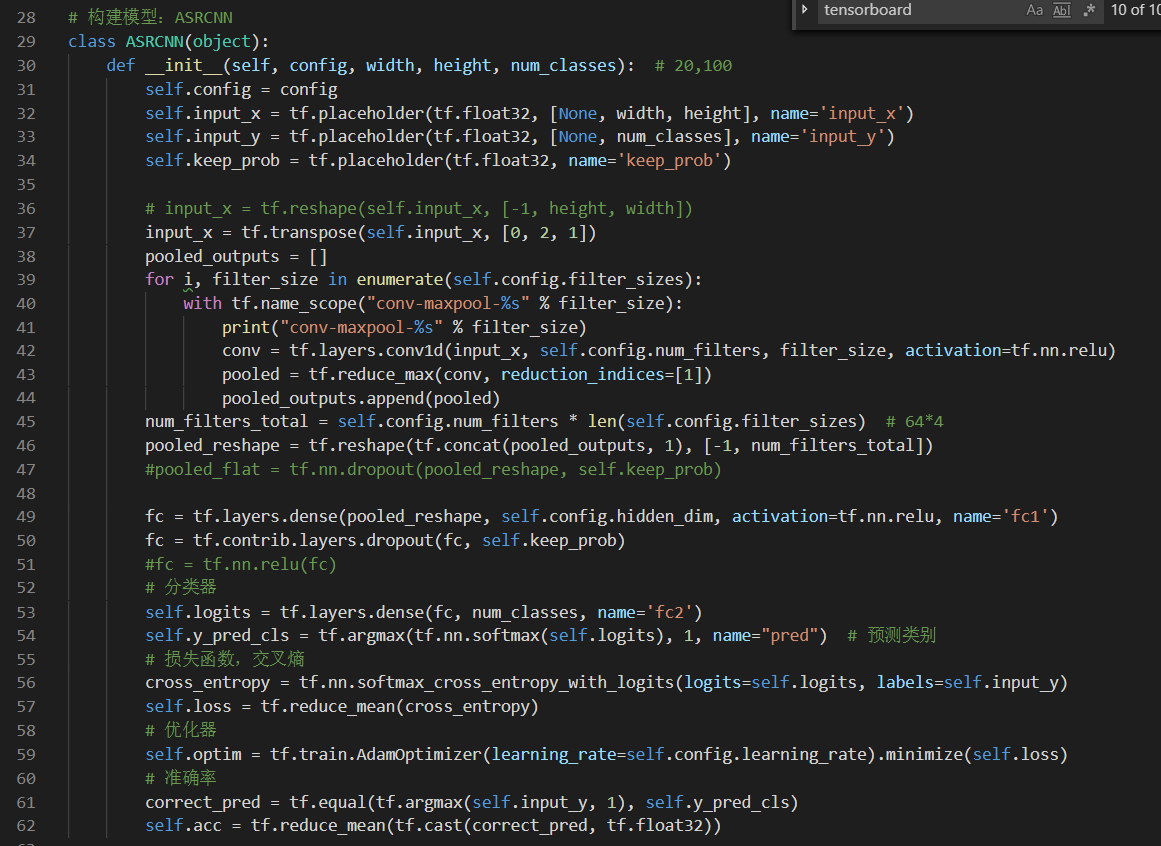
（3）对数据进行批处理



1. 构建模型：ASRCNN

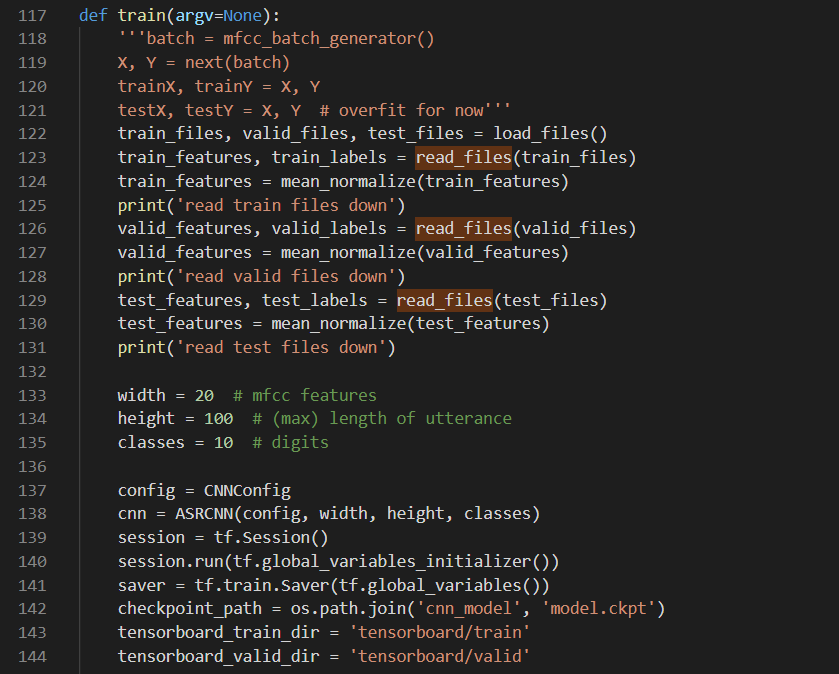
本模型和传统的CNN模型基本一致，只在超参的设置及输入的尺寸存在不同。主要包括：卷积层、池化层、全连接层等。

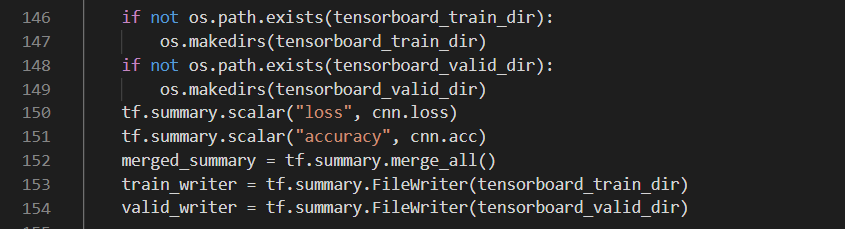


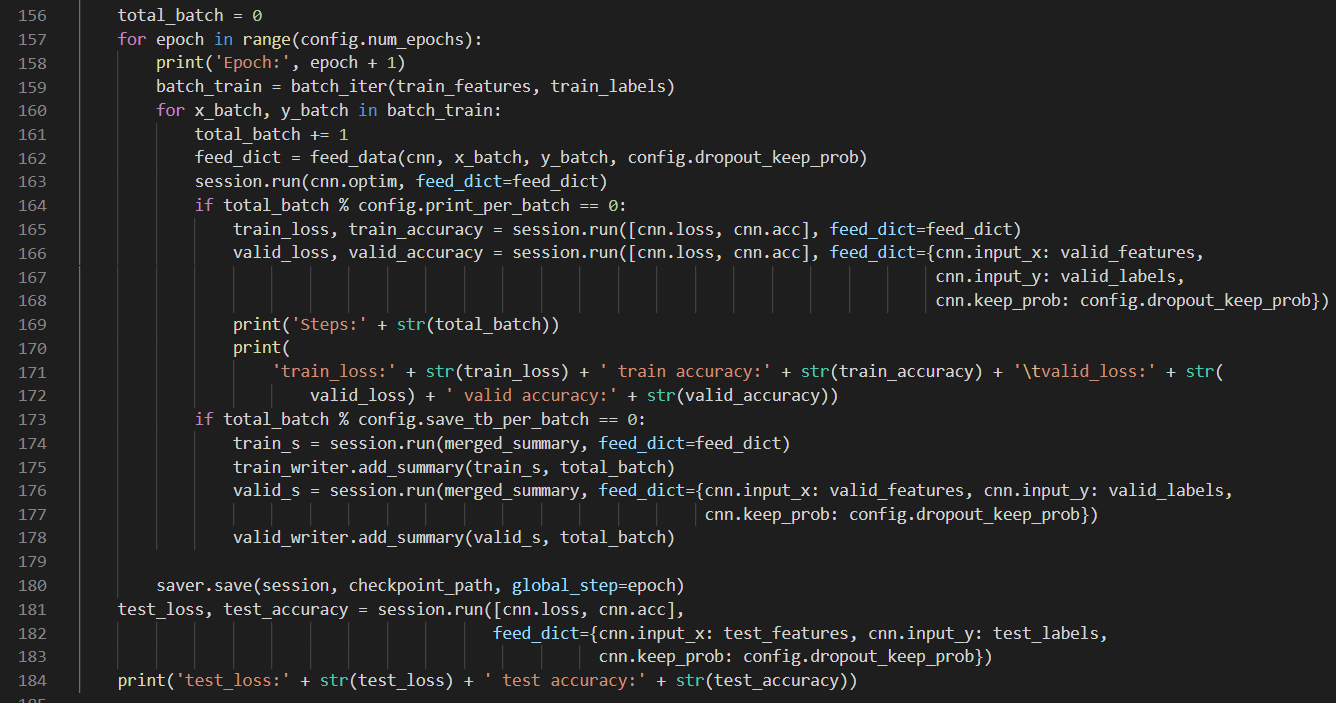


1. 模型训练：

输出模型训练过程中训练集、验证集的损失和准确率值的变化情况，最后输出测试集的指标值，同时保留最终训练结果。



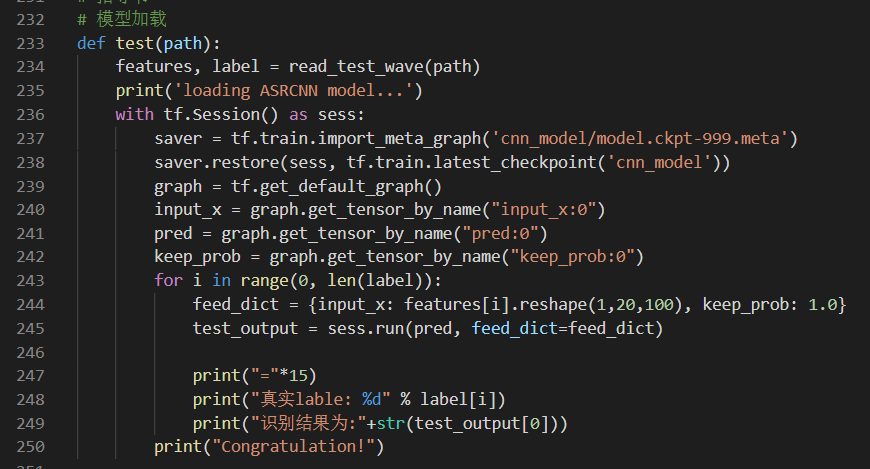




1. 模型使用

加载模型，使用测试数据，输出音频识别结果。



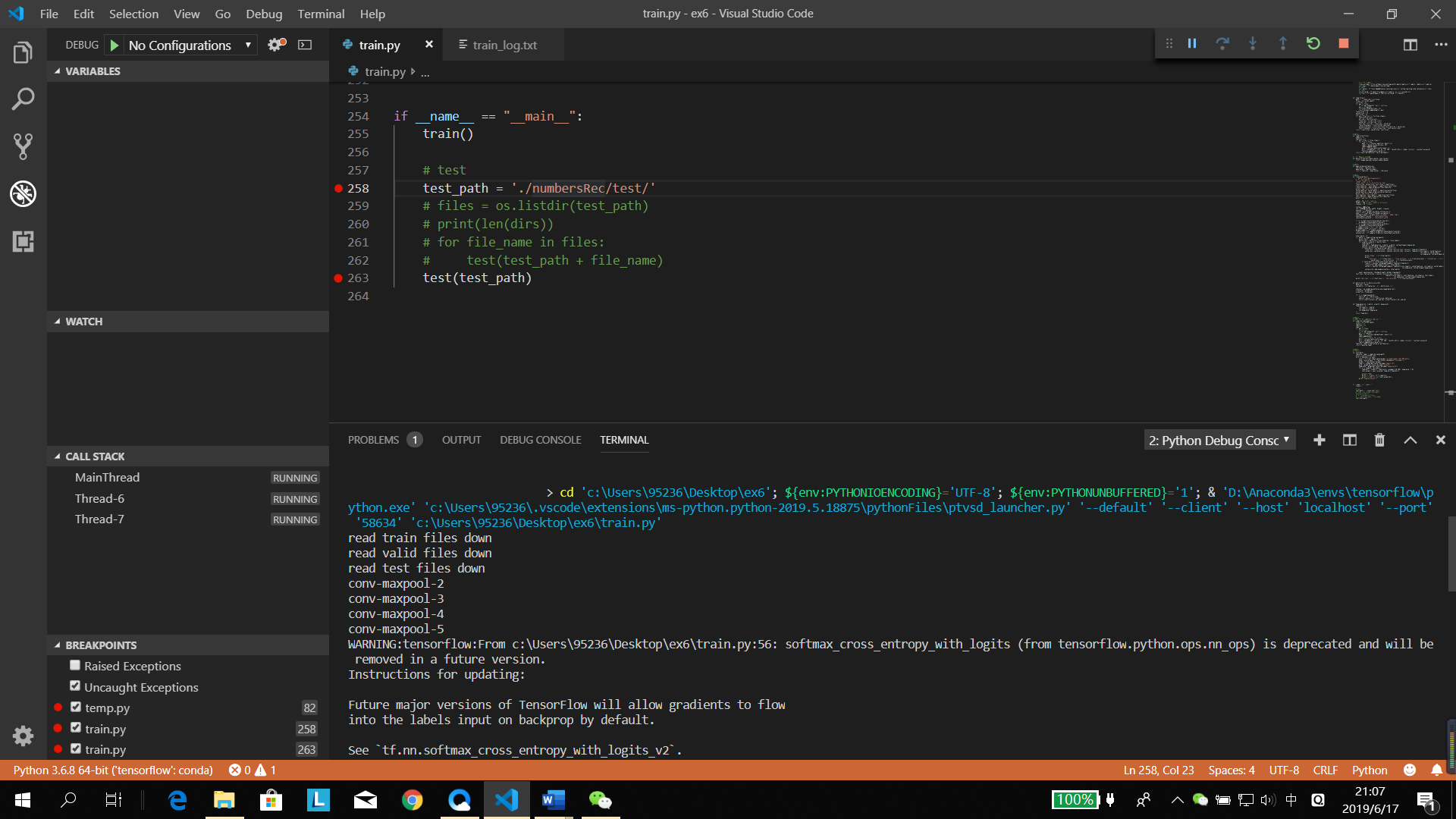


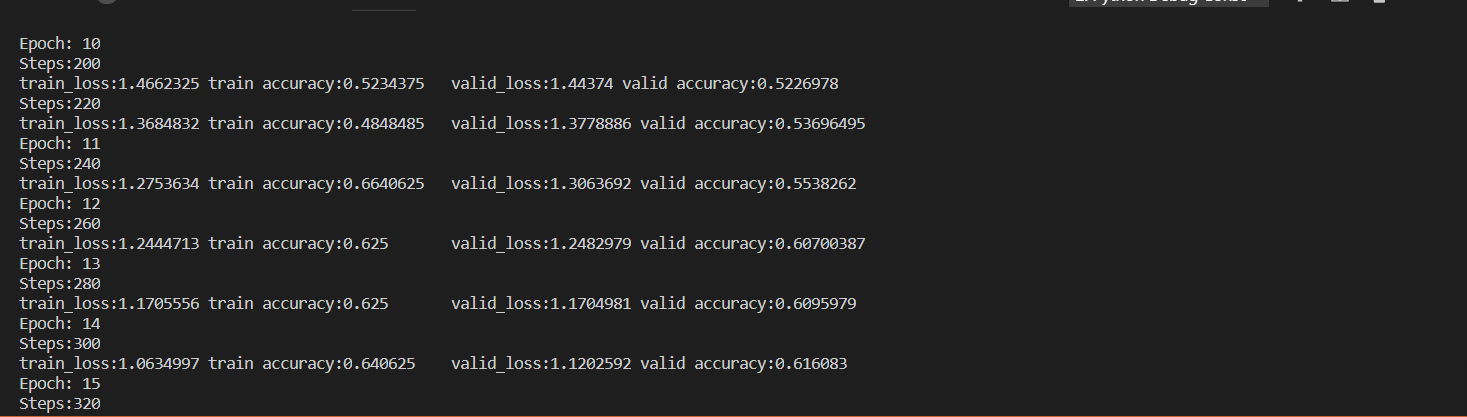
**实验结果：**

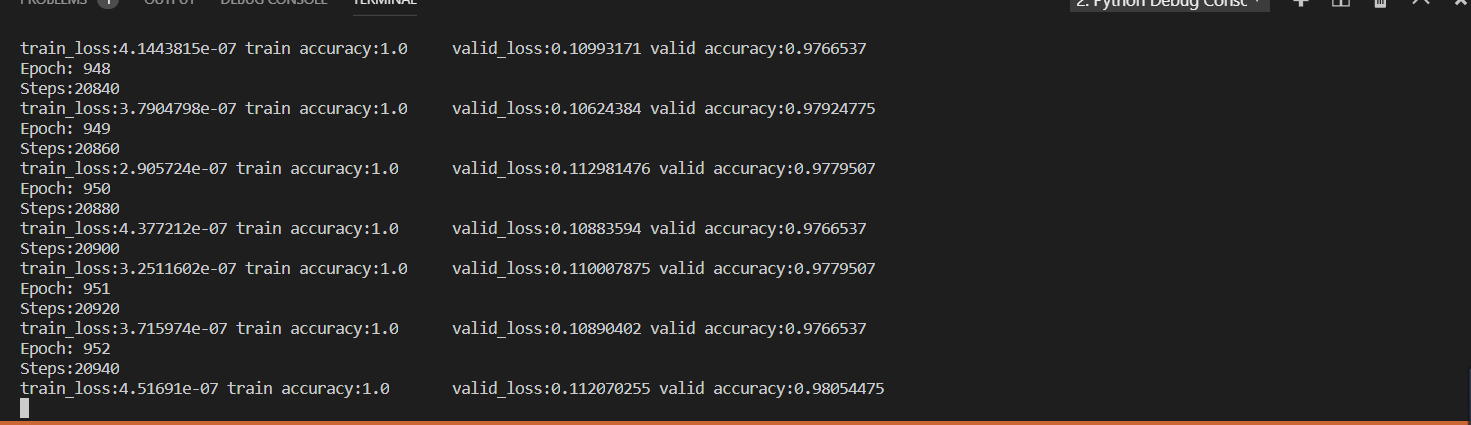
训练和测试分别在train.py和test.py文件里，但是因为为了运行方便且代码量较小，所以当时运行时我把训练和测试同时放在train.py里，只运行一个脚本。

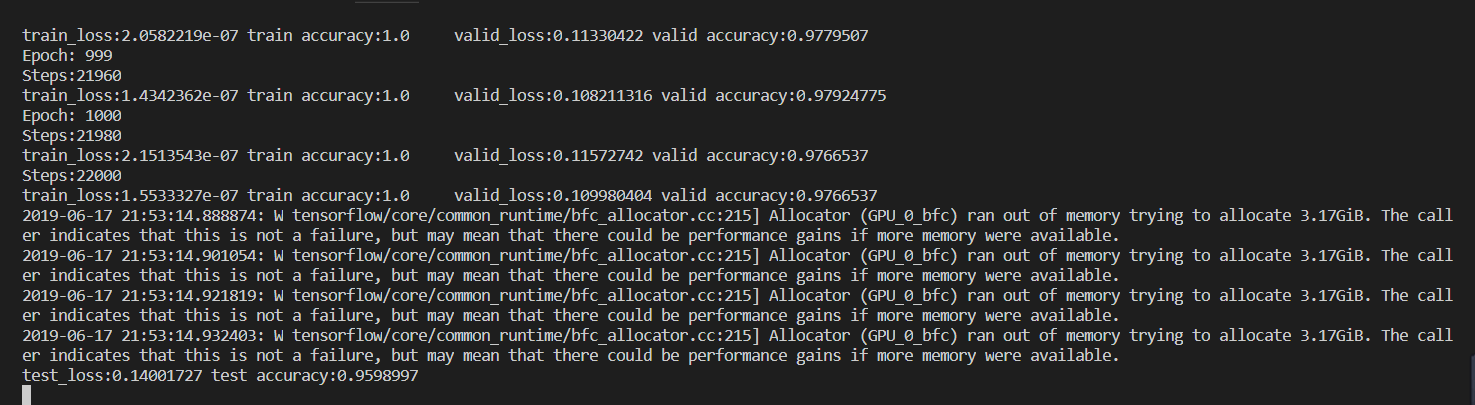
实验开始：

训练：

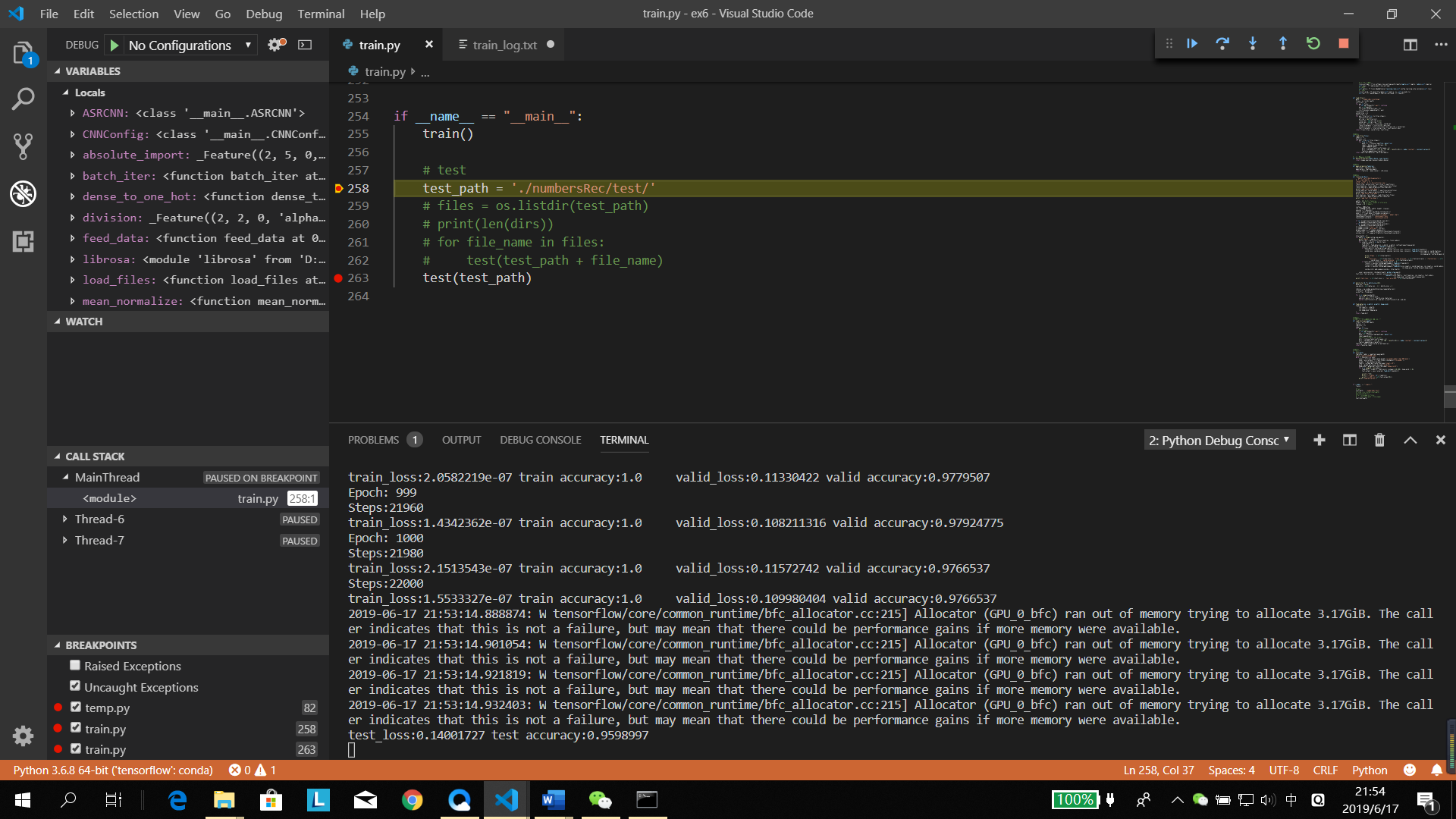


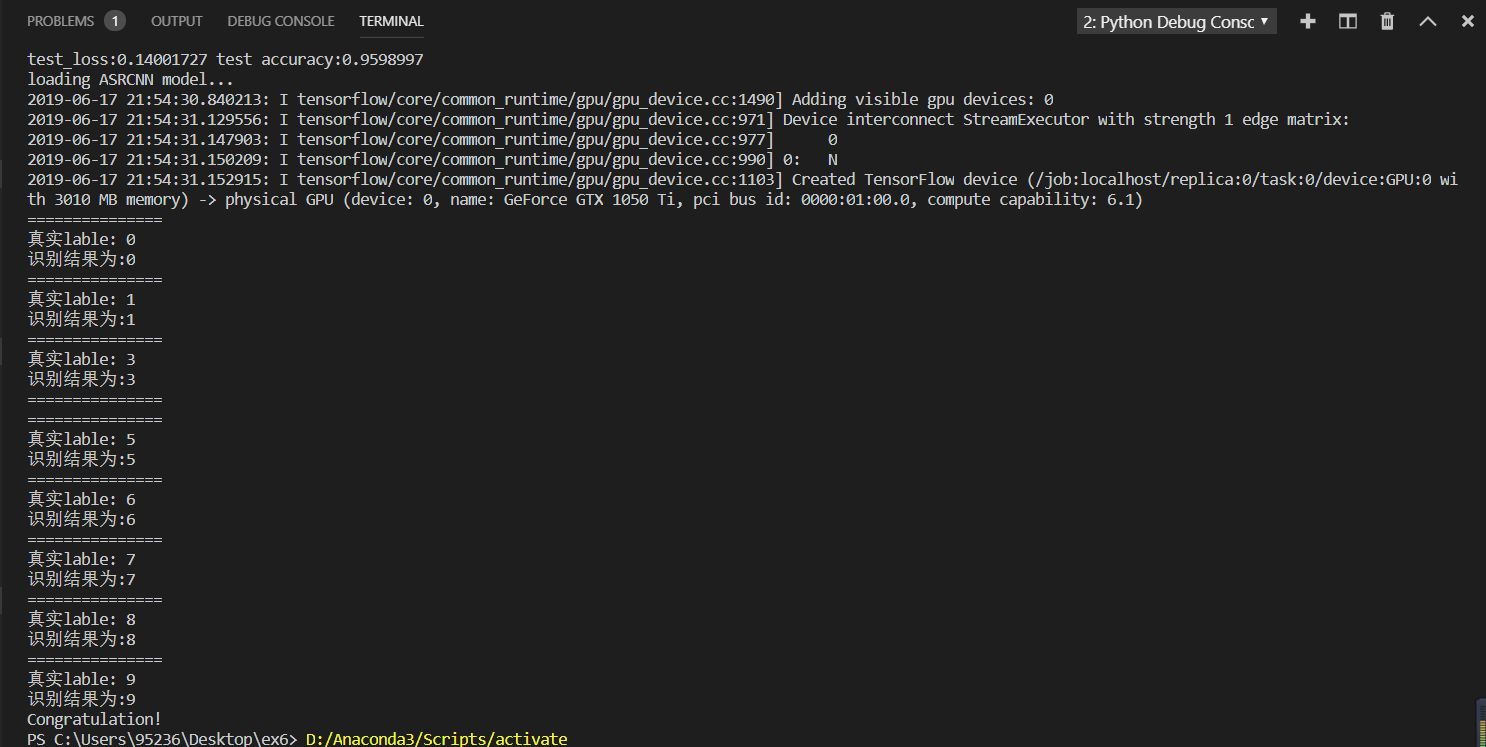






测试：





训练结果图像：

