**土壤分类**

**22360619刘容川**

PS1：本来想先用准确率测测效果的，后面用召回率啥的再跑一遍，结果后面尝试的多了以后就给忘了，想起来的时候确实来不及把所有实验再跑一遍了，所以就用准确率来做评价指标了。

PS2：数据集太少了，划分出来的测试集也太少了，感觉拿所有数据来做测试的话结果会更加可信（我后面用裁剪后语义信息更丰富的图片来训练和测试反而效果更差了，这恰好证明了这个结论）。所以在该任务中我更倾向于推荐使用第二种方法。测试集太少了，后面自己写了个新的数据增强方法，理论推的这种方法效果不错，结果跑出来效果还不如普通的数据增强，就不拿出来丢人现眼了。虽然我觉得是因为测试集太少的原因。

目前最好的结果是最后一种方法中lstm（因为最后一种方法本心是想验证我写的一种新的数据增强方法，所以就随便选了一种轻量化的网络，甚至连CNN都没选）+200epoch，准确率94%，换一种图片裁剪大小后为91%；第二好的是mobilevit训练了10个epoch的，准确率为93.75%；第三好的是clip+prompt（无训练，在整个数据集上测试的，结果更可信），准确率90.62%。因为第一种方法主要是想验证我写的数据增强函数，所以数据集划分和其他方法并不统一。

本文所有方法、数据集划分以及数据增广的代码和效果较好的训练模型可在：

<https://github.com/parkourlrc/prompt-is-all-your-need.git中访问。如需查看请，git> clone。

（clip的模型直接通过代码导入的预训练模型"openai/clip-vit-large-patch14"）

**数据处理：**

1、数据集划分：

抽了0.8做训练、0.1验证、0.1测试。划分的有些不合理。但是前两种方法不用做划分，直接用完整数据集测试就行。

2、数据增广：

下面除了方法6的所有方法中的数据增强均是通过对训练集实行旋转、镜像、亮度变化、裁剪及这四种方式的各种组合操作，实现了对训练集的数据增强操作。

**具体方法：**

1. **基于CLIP的zero-shot分类器**：准确率50%

仅仅使用CLIP和土壤名称来做语义图像匹配，无需任何训练，模型的准确率为50%。

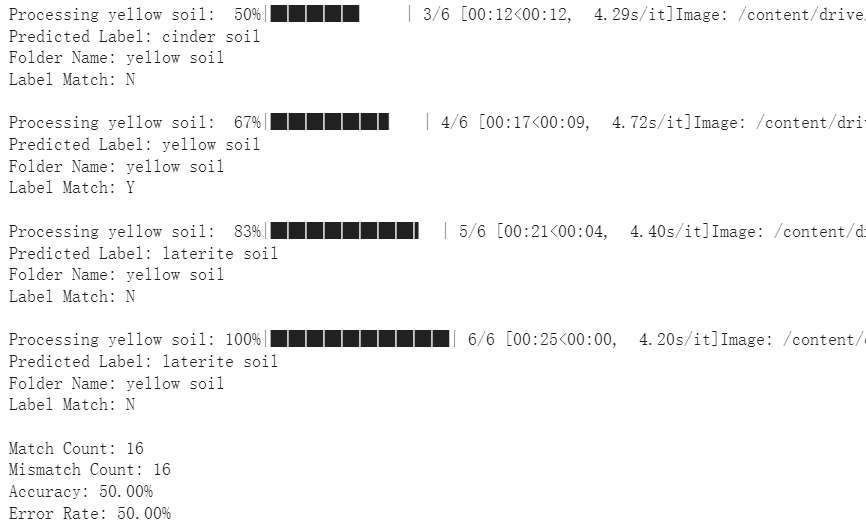
先测试CLIP对土壤分类的有效性：

结果：



然后，直接使用CLIP进行测试（文件结构需要与标签内容相对应）

结果：

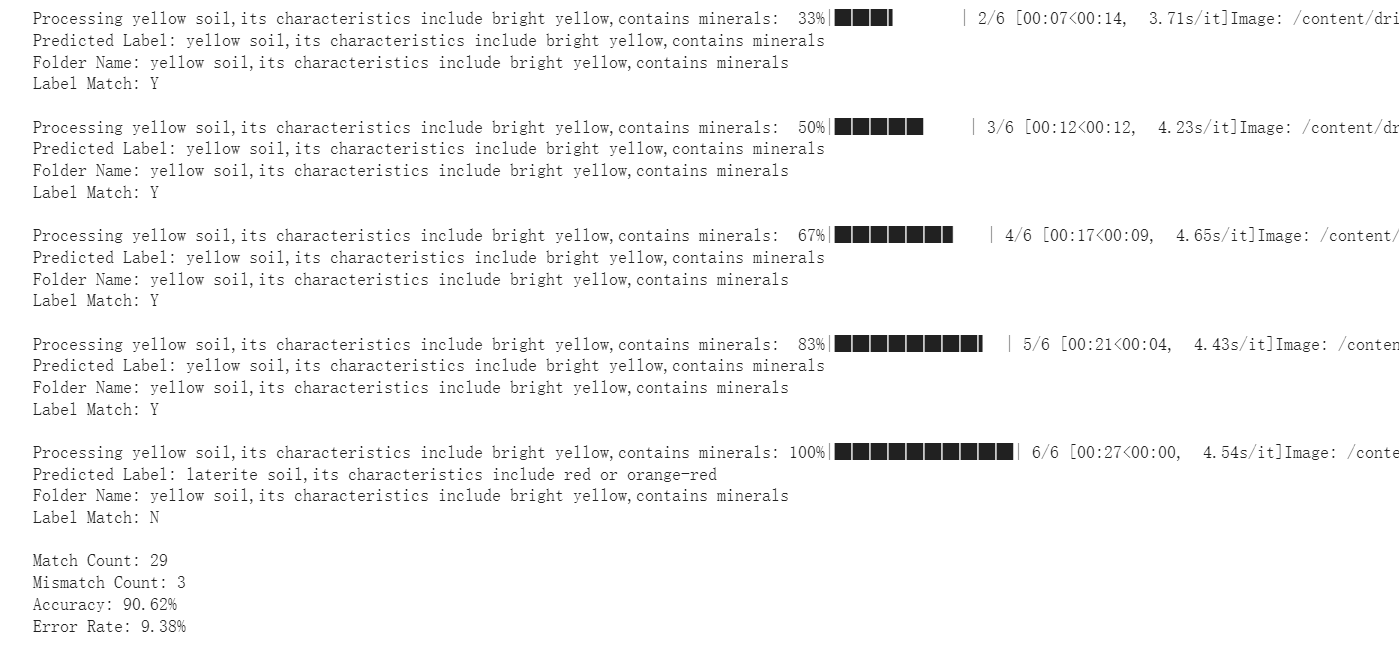


1. **prompt is all your need**——通过仅添加具有语义信息的prompt提高图像—文字空间的匹配效果：准确率90.62%

通过为CLIP所使用的prompt增加具有土壤特点信息的语义信息，模型的准确率飙升到了90.62%。该方法同样也不需要额外的模型训练。

我认为上面这两种方法才是大模型时代的做东西的范式，训练模型对我这种没卡的人来说吃力不讨好（主要是没卡也练不了啥太好的模型）

结果：



1. **小样本微调的对比学习预训练图文大模型方法：**使用训练集数据对CLIP进行微调：准确率78%，训练时长6分钟

没有添加额外的语义信息，就是单纯的做了微调（训练了10个epoch）

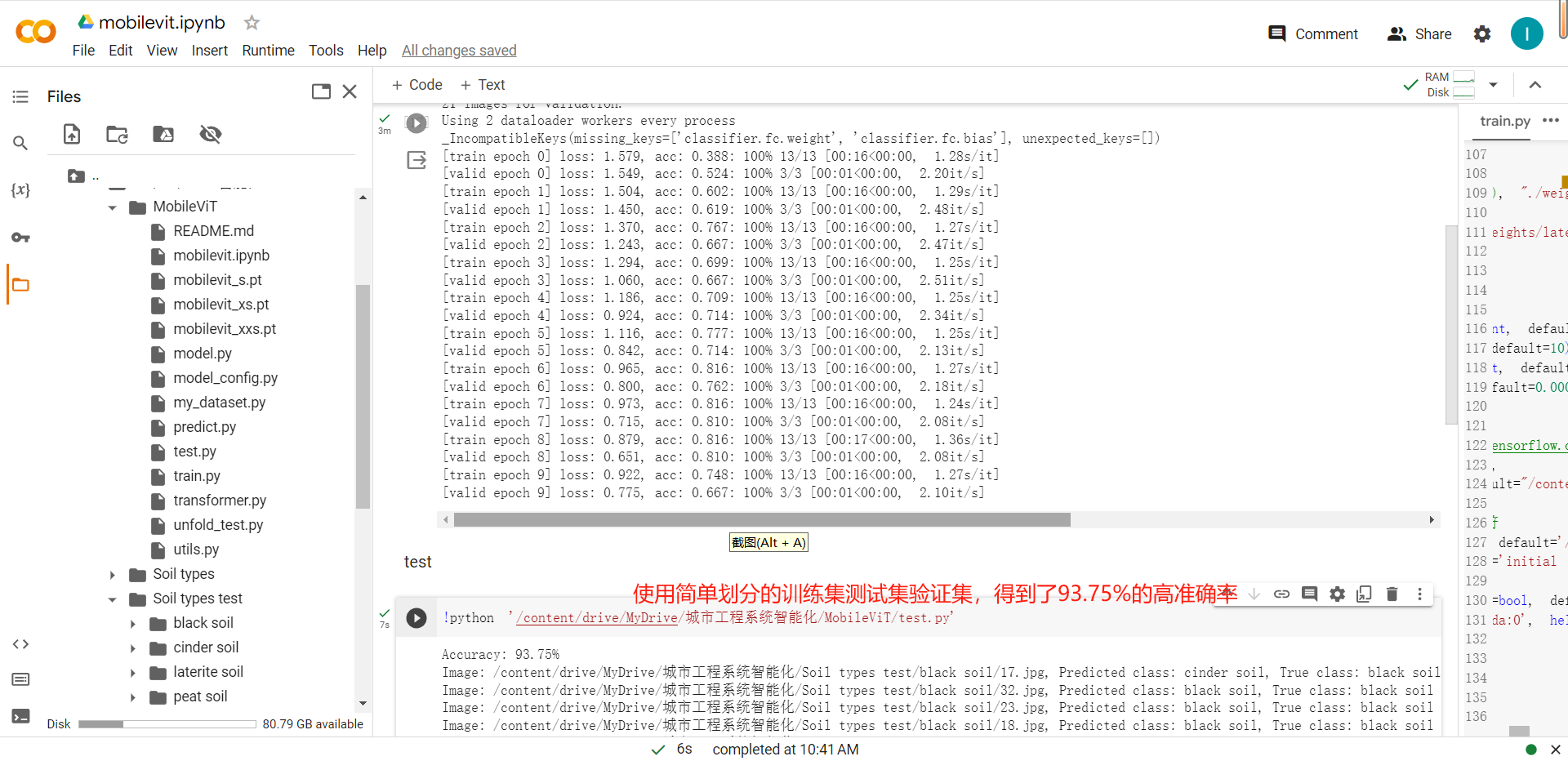
时间太久远了，没截图，这个月GPU限额也用完了，暂时跑不了了，但是效果也不是最好的。

4、**基于数据增强和对比学习预训练图文大模型的小样本土类识别**：使用数据增强后的训练集对CLIP进行微调：准确率90.32%，训练时长3小时30分钟

没有添加额外的语义信息，就是单纯的做了微调（训练了10个epoch）

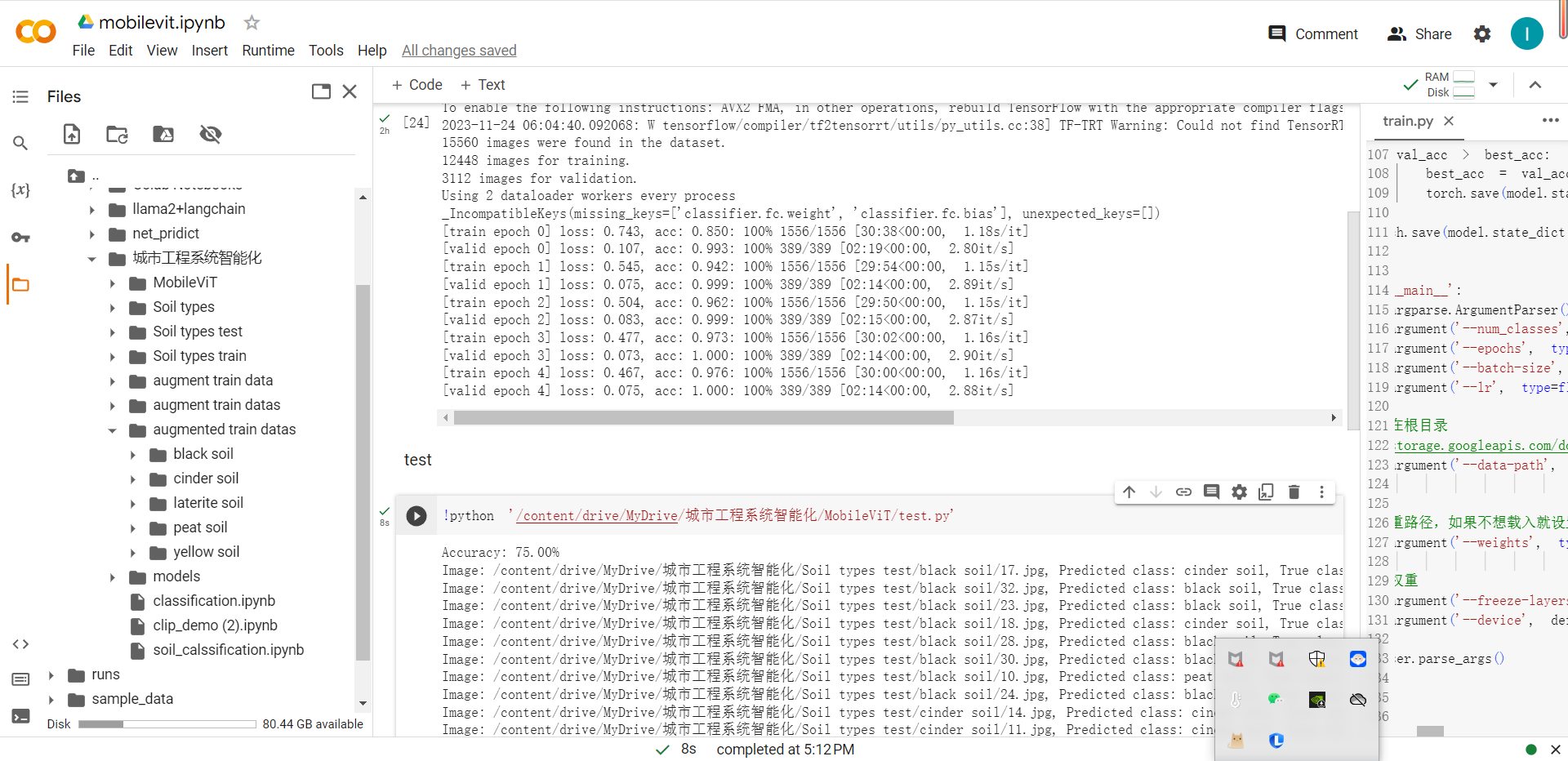
时间太久远了，没截图，这个月GPU限额也用完了，暂时跑不了了，但是效果也不是最好的。

1. **基于MobileViT的小样本土类识别：**使用训练集对MobileViT进行训练（5个epoch）：准确率68.75%，训练时长1分半。当训练10个epoch时，准确率为93.75%，训练时长3分钟。



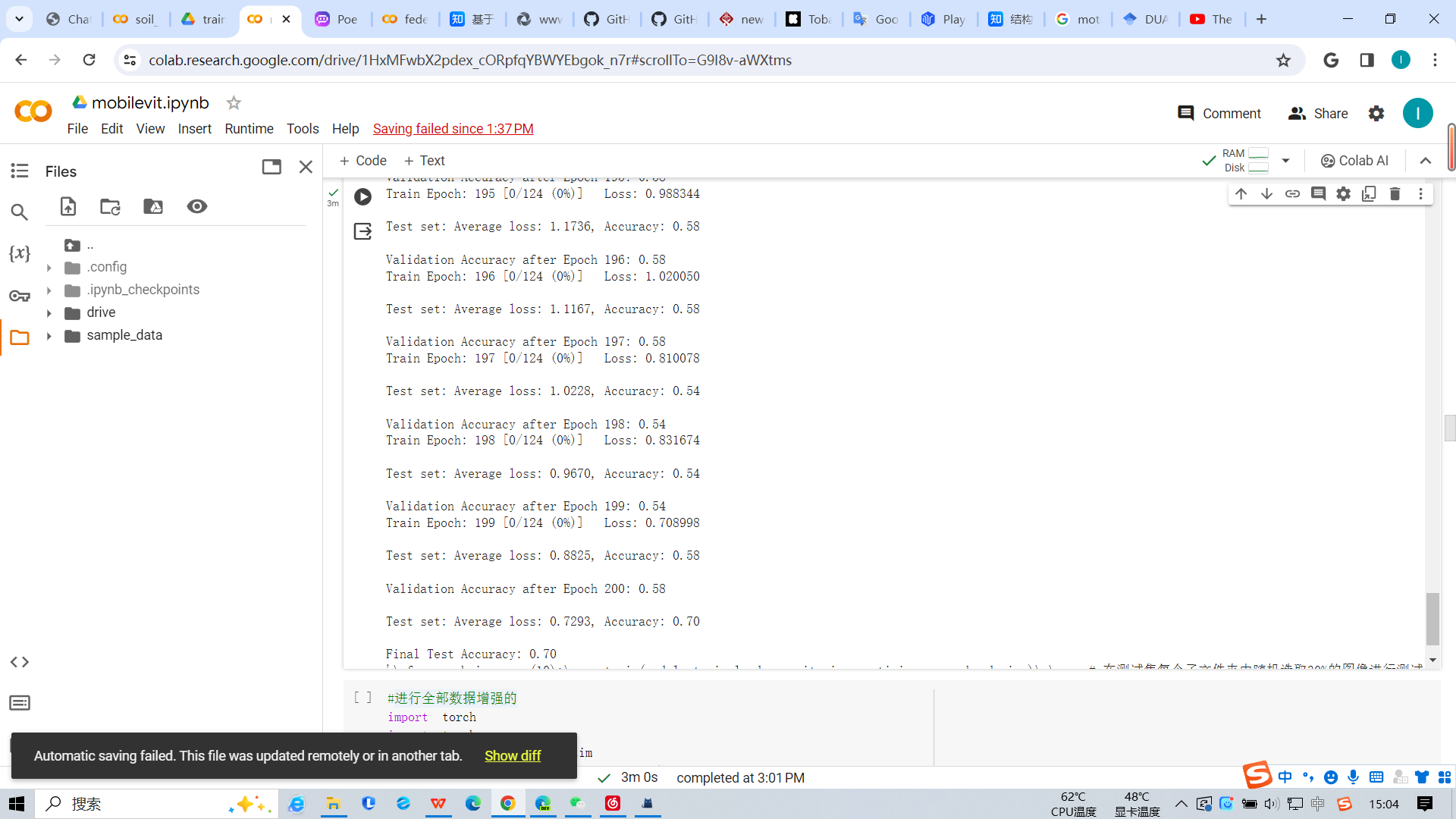
1. **基于数据增强与MobileViT的小样本土类识别：**使用数据增强后的训练集对MobileViT进行训练：准确率75%，训练时长2小时（由于时间太长的原因只训练了5个epoch）

结果：

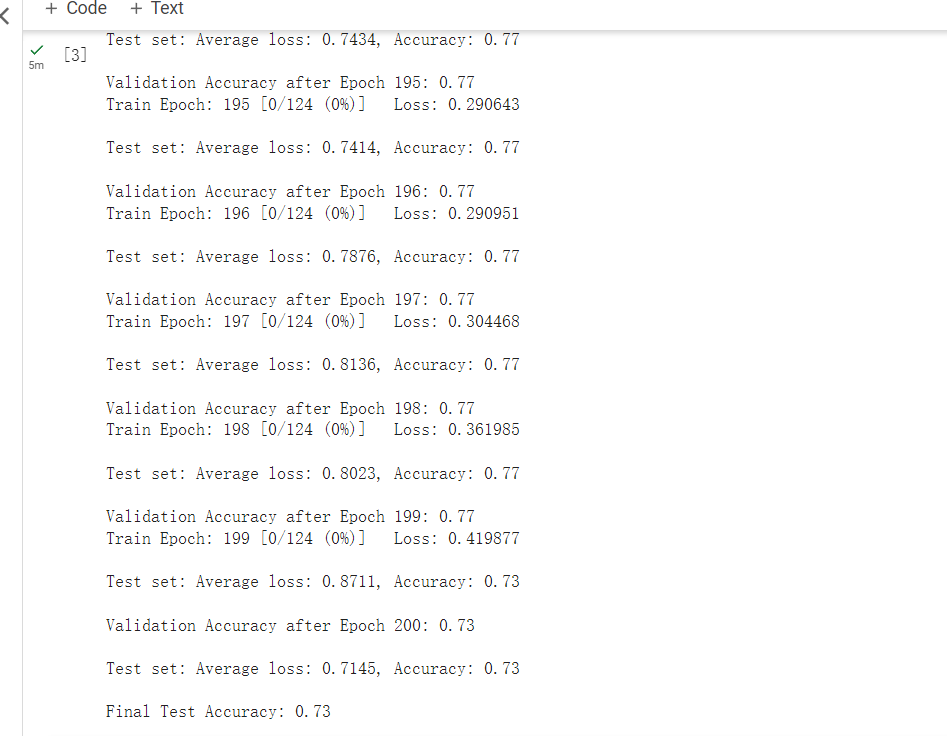


1. 基于自适应增强比的数据增强方法的在土壤分类上的应用

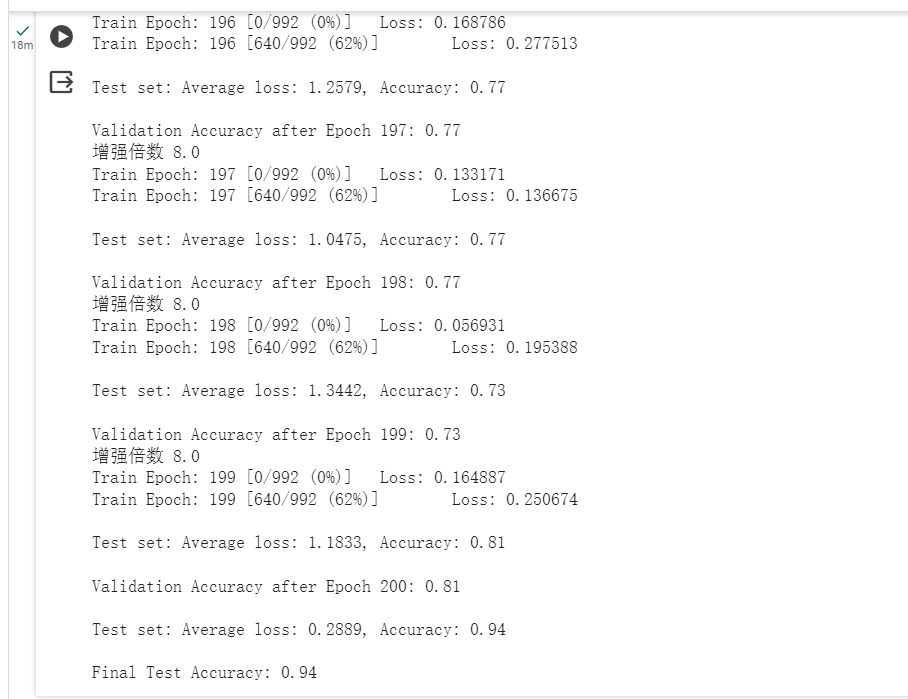
无数据增强的lstm，图片大小64\*64



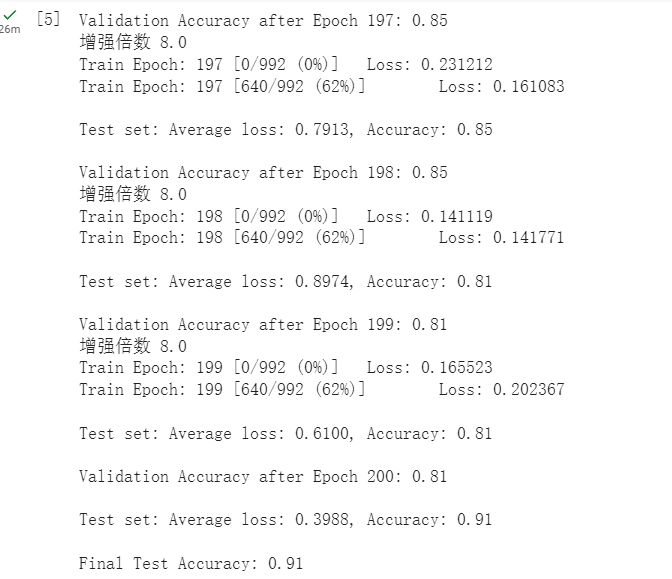
无数据增强的lstm，图片大小128\*128：



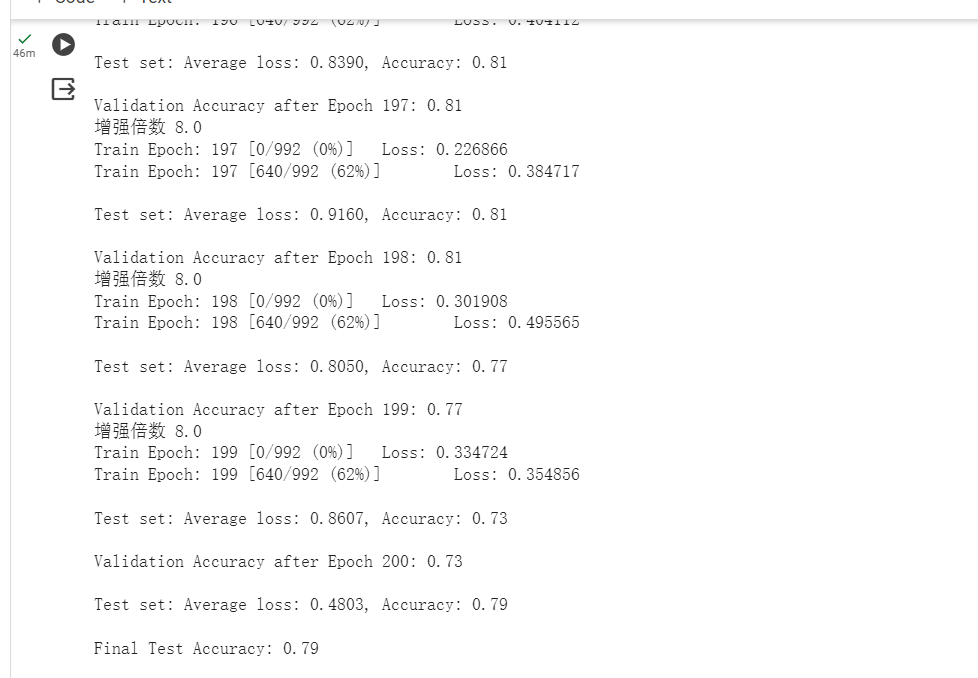
**全部数据增强的lstm，图片大小32\*32（效果最好94%）：**



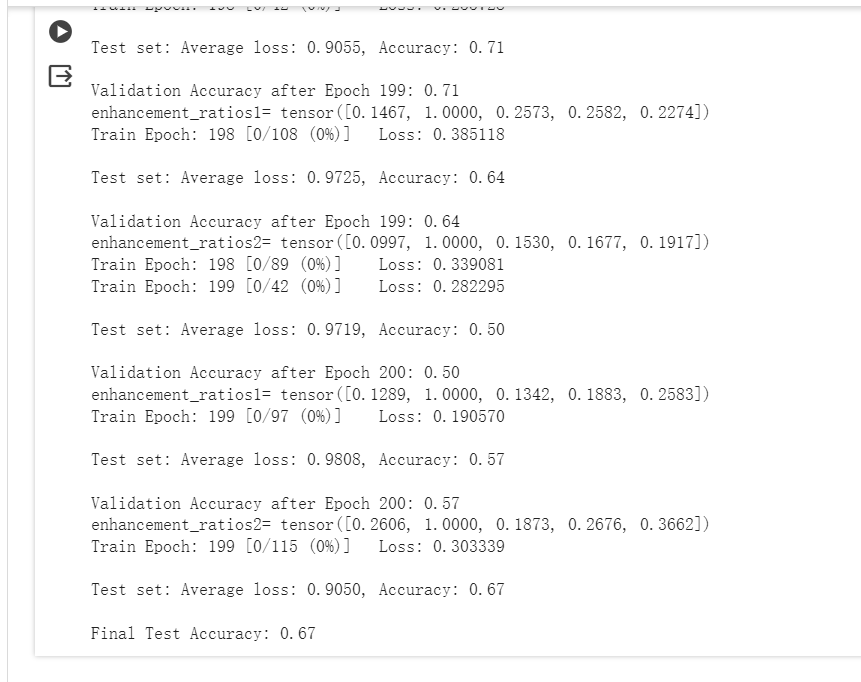
全部数据增强的lstm，图片大小64\*64：



全部数据增强的lstm，图片大小128\*128：



基于自适应增强比的数据增强的lstm，图片大小32\*32：



基于自适应增强比的数据增强的lstm，图片大小128\*128：

