NLP自然语言大作业

姓名：朱良双

日期：12月20日

|  |  |
| --- | --- |
| 硬件环境 | GPU NVIDIA Geforce RTX3060 |
| 操作系统 | Windows11 |
| 采用的深度学习框架  工具  语言 | Pytorch  Pycharm，anaconda  Python |
| 问题描述 | 参考博文<https://blog.csdn.net/dgvv4/article/details/125184340>实现对动物的三分类，调整超参，实现对花卉的五分类。 |
| 对模型的改进 | 将论文Zhou D , Kang B , Jin X , et al. DeepViT: Towards Deeper Vision Transformer[J]. 2021.所提的Re-attention模块融入DeepVIT模型，达到对模型预测准确率和速度上的提升。将batchsize设置为8和12和原本的模型进行精度对比。 |

文件讲解：

deep\_VIT.py: 再注意力模块 Re-Attention的VIT模型

draw\_picture.py:绘制loss和Acc曲线

VITflowerClassification.py: vit模型训练鲜花5分类

DeepVITflowerClassification.py：采用deepVIT模型鲜花5分类

Image\_standAndmean.py: 数据预处理得到图片的标准和均值

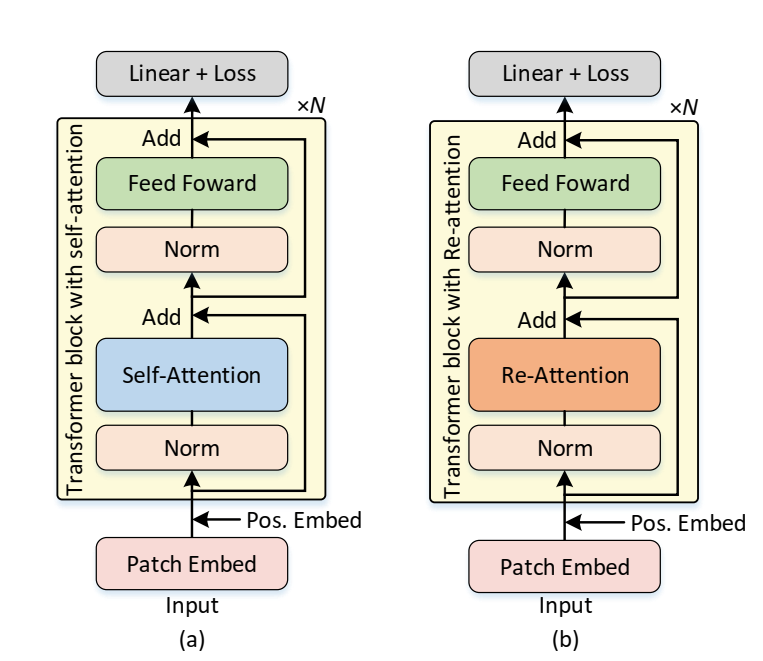
VITpredictFlower.py：VIT鲜花分类测试

DeepVITpredictFlower.py：DeepVIT鲜花分类测试

VisionTransformer\_model.py: 传统VIT模型框架

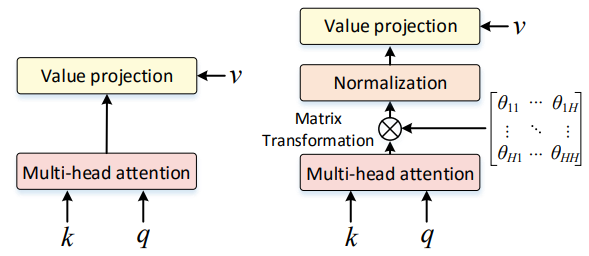
模型介绍

经典的 Transformer 由 Encoder 和 Decoder 组成，其中，最重要的就是多头注意力机制（Multi-head attention）。在 DeepViT: Towards Deeper Vision Transformer 论文中，作者将经典的self-attention模块替换为Re-Attention，能够考虑到不同attention之间的信息交换，模型框架如（b）所示。

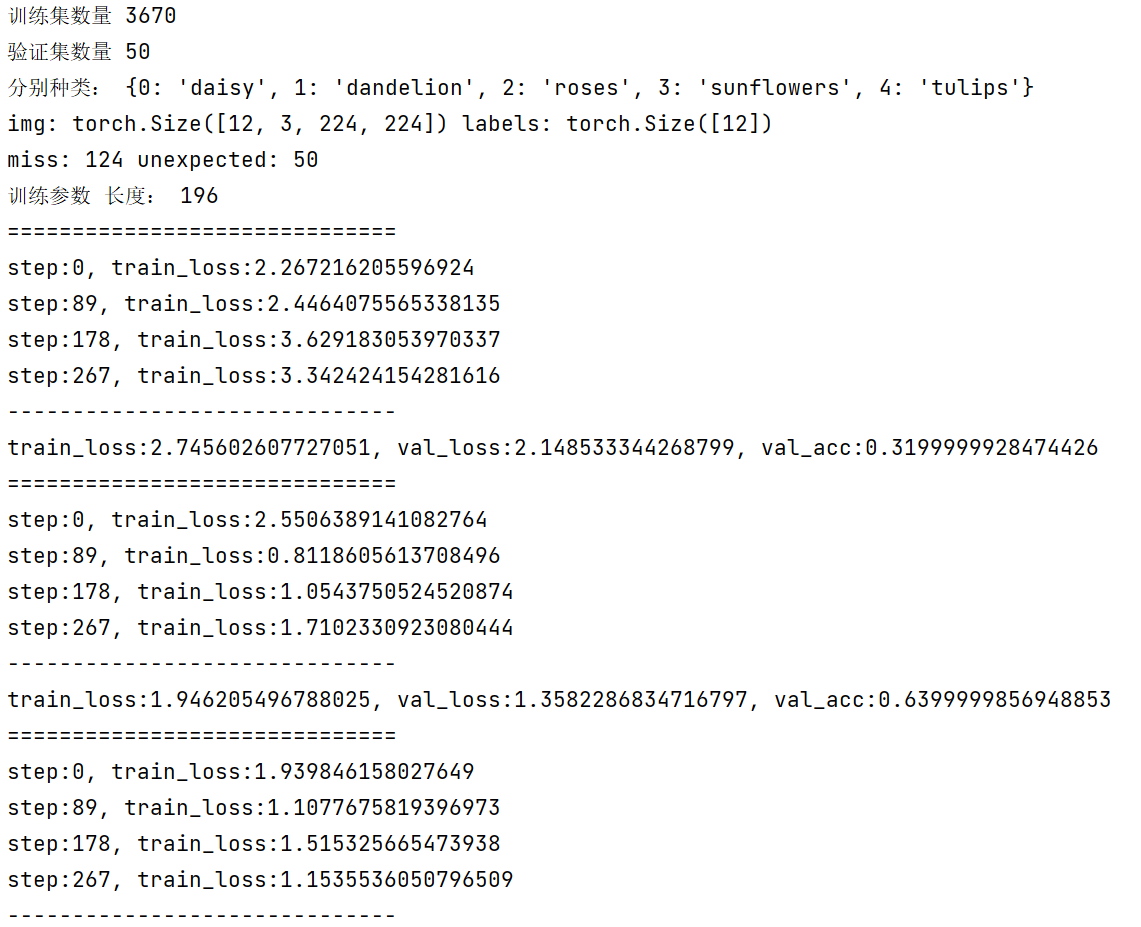


，

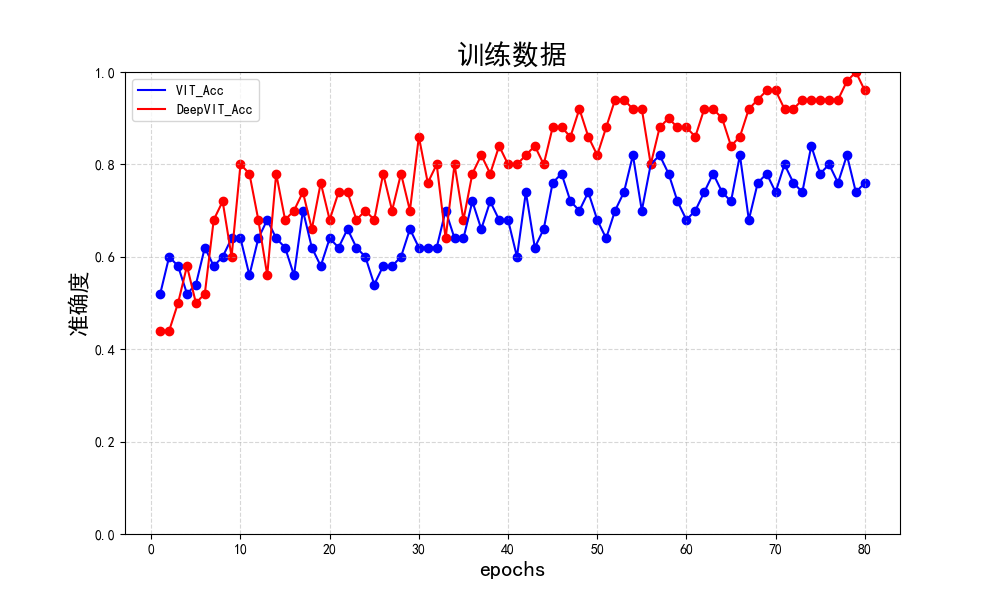
在Multi-head attention后加入一个可学习的权重矩阵，实现再注意力机制，如下图所表示。

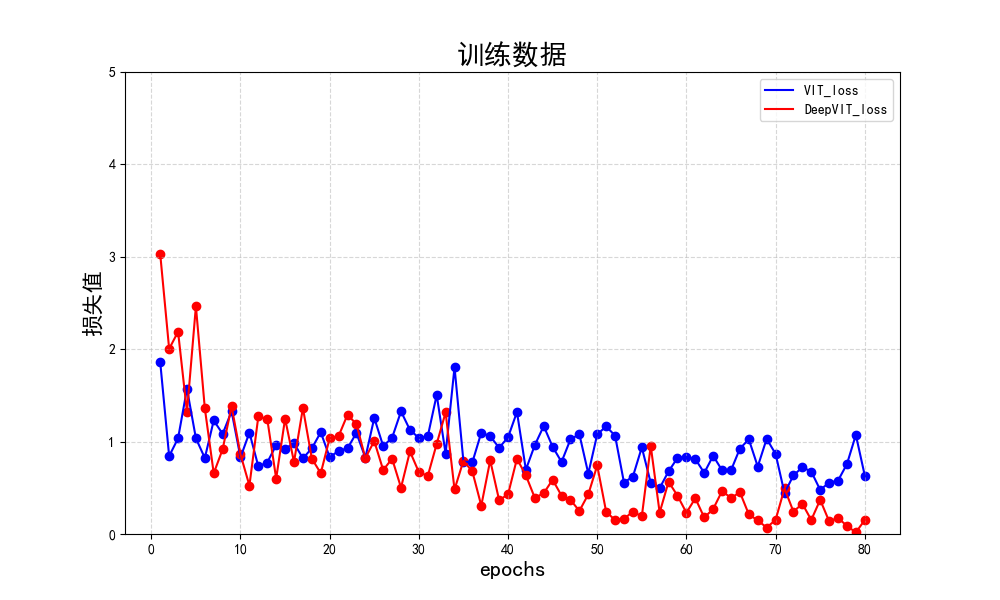


程序运行：运行DeepVITflowerClassification.py和VITflowerClassification.py文件训练模型得到训练完成的参数文件，训练过程如下图所示

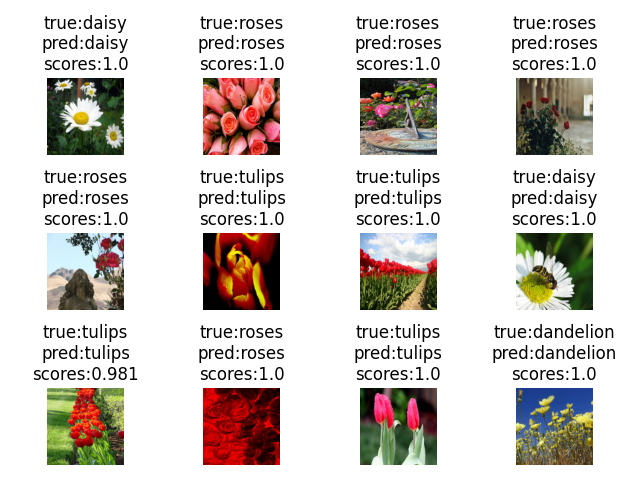


模型Acc与loss的值对比如下，经典VIT模型训练的Acc值要比DeepVIT模型





测试DeepVIT运行结果如下图



鲜花数据集链接: https://pan.baidu.com/s/1OJfwcF1PvX9qkZwT7MXd\_Q?pwd=i0bt 提取码: i0bt