|  |  |  |
| --- | --- | --- |
| SRT计划项目个人研究报告 | | |
|  | 项目名称: | 基于Keras深度学习框架的多标签图像分类研究 |
|  | 申请者: | 邱日 |
|  | 学 院: | 信息科学技术学院 |
|  | 专 业: | 计算机科学与技术 |
|  | 指导教师: | 顾兴健 职称: 讲师 |
| 2018年 5月 11 日 | | |

首先感谢一年多来耐心指导我们小组的顾兴健老师，也感谢我们SRT小组的两位组员秦天涯和周宽．

一年半前，有幸请到顾老师来做我们SRT小组的指导老师，从此投入到了对深度学习的探索中去.我作为组长，也组织了积极上进的两位同学秦天涯和周宽来与我合作展开SRT的研究．

顾老师根据自己的研究方向，给我们确定了＂多标签图像分类＂的研究课题.在技术路线上决定使用Python作为开发语言；在深度学习框架上，当时有几种可供选择：一是Caffe,二是theano及建立在theano之上的Keras．

机器学习研究历史比较长，但是深度学习的最近一次崛起还是在前几年，Caffe虽然到2013年才被加州大学伯克利分校的贾扬清博士建立起来，但在业界已经算是深度学习框架的＂元老＂，然而配置麻烦，扩展起来不灵活；Keras

始于2015年，才刚刚在社区出现，不过使用起来相对简单，模型建立不需要单独的配置文件，对新手相对友好，利于研究者比较快地将好的想法转化成代码．

当时Keras深度学习框架还算是非常小众的，据我们所知，中国此时没有一本详细讲解Keras框架的出版图书，国外除了Keras官网相关资料也很少．顾老师和我们商讨之后确定使用Keras,后来的事实证明这是十分有远见的，之后Keras被谷歌热门的深度学习框架TensorFlow所看中并合并，被越来越多的研究者所青睐．

深度学习研究者普遍在Linux下开展研究，Ubuntu操作系统是个合适的选择．通过顾老师，我们第一次实际接触Linux.之前我们都一直使用Windows,对Linux安装和使用没有概念，我作为组长，积极利用图书馆和网络搜索相关资料，

并结合自身不懈的实践，初步掌握了Linux命令的使用，对环境的配置也有了了解，先把自己的实验环境配置好，再帮助小组成员部署环境，耐心讲解命令的使用，小组成员Linux的使用水平也有了很大提高．

SRT提高了我们的综合能力，我们学会熟练使用Github找到有意思并值得研究的相关项目，学会在Kaggle数据分析和机器学习网站上在线编程并在云上训练．寻找一个合适的数据集是个很重要的问题，可用于多标签图像分类的数据集倒不少，不过绝大多数都非常的大，不适用于我们的情况．顾老师建议我们用ImageNet数据集，我们就花了一周的时间将几百GB的ImageNet数据集下载了下来，在初步探索之后发现所需要的计算资源是我们所不能承受的．我们向顾老师反映了这个问题，之后的建议是选择其中的五类来训练．

我们通过研究发现，ImageNet图像多为单标签，当数据量比较小时并不适合我们的多标签图像分类任务．之后我们找到了南大周志华教授公开的多实例多标签图像数据集，这个数据集专门用于多标签图像分类，图像标签分为五类，

有＂海＂、＂树＂、＂日落＂、＂沙漠＂、＂山＂．数据集大小对我们也正合适，总共24MB,一共有2000张自然风景图像，还有.mat格式的图像标签．

在模型的建立上，我与秦天涯一起探讨预训练模型的’fine-tuning’和新层的添加；周宽广泛阅读了一些论文，理解其中的概念，和我们探讨，在数据集的选取上做出贡献，也帮助我们解决了很多技术上的问题．我作为组长，编写了大部分核心代码，周宽和秦天涯研究了TFrecord的格式，解决了数据预处理的问题，也和我探讨了预训练模型的选取和卷积、池化、全连接层等的作用，大家确定了在multi-lable问题中使用sigmoid函数替代softmax函数作为激活函数．由于我和周宽的电脑内存和性能不满足要求，大家使用秦天涯的电脑完成了模型的训练，最后可以达到百分之九十五以上的正确识别率．

通过SRT，我们的理论基础和动手能力都得到很大提高．我作为组长，有效地组织了大家的学习活动，同时帮助了省级国家级和校级SRT做相关研究的几个小组同学配置环境和解决问题，也从他们那里学到不少东西．通过合作交流，沟通协作能力也得到锻炼．唯一感到遗憾的是现有实验条件使我们研究的内容受到限制，如果训练模型时能够得到学校服务器的支持，相信可以做出更加优秀的作品出来．