|  |  |  |
| --- | --- | --- |
| 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: | 秦天涯 |
|  | 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 | | |

1. 项目介绍

随着照片和视频分享网站的普及（Flicker,Instagram 和 Youtub等等），大量的未标记的和弱标记的图像数据也随着传播开来。这越来越多的数据图像需要有效的数据检索来达到获取这个图像的最终目的。尽管图像检索已经被研究许多年了，搜索引擎检索目的图像仍然大多基于文本内容而非原始图像语义。因此，自动将图像标注成人类已经提供好的关键词或者标签的技术已经受到了大量的研究关注。该技术从开始到现在的发展大致经历了三个阶段，分别是：基于文本的图像检索，基于内容的图像检索，基于语义的图像检索。基于文本的图像检索，由于手工标注工作量大，个人理解也有差异，导致检索结果出错率较高；基于内容的图像检索，是根据图像的底层特征进行检索，没有反映图像的内容信息；而基于语义的图像检索，弥补了上述两种方法的缺陷，使检索结果和使用者所需要的信息尽可能的一致。对于语义标注可以人工标注或者机器标注，由于上述所讲个人的理解因人而异，所以现在大多使用机器自动标注，即将等待标注的图像输入计算机系统，经过分割，将它划为等大的方格，而每个小方格的特征都被一个视觉特征词代替，让系统通过一系列算法推理出可能划分的语义，并将整个图像转为语义序列，相邻而重复的视觉词去掉然后将每个视觉词标注它所对应的物体标签。

2.研究目的与方向

本次SRT我们的研究课题是：基于 Keras 深度学习框架的多标签图像分类研究。是以多标签图像标注为目的，在Keras深度学习框架下，使用Python语言，结合TensorFlow库，开发图像标注的核心算法。结合ResNet50预训练模型和以上算法训练出属于我们自己的模型，在预训练模型的基础上提高其泛化能力及其图片识别的精准率，最后实现对图片的多标签分类。这就是我们课题的研究方向。

3.开发工具以及研究手段

3.1研究环境以及工具

**Ubuntu系统**

**基本的python语法**

**TensorFlow** TensorFlow 是谷歌基于 DistBelief 进行研发的第二代人工智能学习系统，其命名来源于本身的运行原理。 Tensor（张量）意味着 N 维数组， Flow（流）意味着基于数据流图的计算， TensorFlow 为张量从流图的一端流动到另一端计算过程。TensorFlow 是将复杂的数据结构传输至人工智能神经网中进行分析和处理过程的系统 TensorFlow 可被用于语音识别或图像识别等多项机器学习和深度学习领域，且完全开源，任何人都可以用。

**Keras** Keras 是一个高层神经网络 API， Keras 由纯 Python 编写而成并基 Tensorflow、 Theano 以及 CNTK 后端。 Keras 为支持快速实验而生，能够把你的 idea 迅速转  
换为结果

3.2研究手段

**Kaggle** kaggle是一个与我们SRT密切相关的工具，他可以实现在云上编程，我们的代码需要一些基础的环境和一些数据集，这个网站上会为我们提供这些环境和数据集（数据集都是一些比价小的打的跑起来比较麻烦），在这个网站上还有一些专业人士提供的相关代码我们可以参考参考。

**GiHub** github乃程序员的三宝之一，这个网站是我们常用的网站之一，上面提供了相当专业的代码，在这个网站搜索关键字能最快的到我们想要的东西，上面的每一行代码都有着英文注释，这对我们SRT的进行有了极大的帮助。

4.研究学习过程

4.1 基本环境的安装

工欲善其事，必先利其器。我们的第一步是安装**Ubuntu**系统，以及配置python编程环境和keras以及tensorflow函数库。有了以上编程环境，SRT才能继续进行。

4.2基本语言以及语法的掌握

由于本次要用的库是TensorFlow和 Keras，这两者是建立在对python基本语法了解的基础之上的，所以说在SRT 进行的初期，我和组员一起购买了《从零开始学习python》这本书，之后还在组长邱日那里拷贝了python学习视频进行学习python。这就是我们SRT初期做的准备。

4.3 Keras相关术语以及Keras内容的学习

这里我们学习Keras主要通过看Keras中文文档学习，这里的文档由邱日组长提供，上面介绍了如何构造一个模型，以及一些API的介绍，由于我们都是深度学习的初学者，这里我们花了大量的时间来看Keras中文文档。

4.4深度学习相关例子试跑

在有了python基本语法以及Keras的一些基本的理论知识之后，我们的导师顾老师给我们布置了任务，去相关网站上找到深度学习的代码去试跑，这里有邱日组长提供了几组代码，组长耐心地教导们我如何将这些代码跑起来，遇到了错误也一步一步帮我们解决，接新事物的我们此充满了好奇。

4.5数据集和预训练模型的了解和选取（这里是由组长决定的）

在做了之前的了解之后,我们就要开始构造自己的训练模型了,在构造模型之前我们要选取数据集和预训练模型,由于我们组做的是多标签图像标签,这里通过相关网站了解到,比较先进和适合我们的预训练模型是resnet50,

数据集选的是miml-image-data，数据集由 2000 张自然风景图片和标签组  
成：图片的内容有山、水、日落、海、树这五类组成，其中每个图片内容包含一类或多类; 图片对应的标签在.mat 格式的文件中给出。Resnet50函数原形如下：

keras.applications.resnet50.ResNet50(include\_top=True, weights=’imagenet’,  
input\_tensor=None, input\_shape=None,  
pooling=None,  
classes=1000)

4.6激函数的选取

我们可用的激活函数由softmax函数和sigmoid函数，由于但标签和多标签的问题我们选取的函数使sigmoid函数，Sigmoid 函数的优点在于，它的输出映射在（0,1）内，单调连续，非常适合用作输出层，并且求导比较容易。但是，它也有缺点，因为软饱和性，一旦落入饱和区，导数就会变得接近于 0，很容易产生梯度消失。

4.7迁移学习与模型的构建

迁移学习的意思是在选取的预训练模型的基础上继续学习，进行模型的改造与加层，这部分主要围绕组长和我来展开，有我来协助组长完成，起初我们先将预训练模型加载进来，直接用预训练模型进行预测，测试预训练模型是否加载成功，之后在根据我们数据集的特点以及参考其他代码的写法完成了模型的构造。（此时的模型能否成功还未知）

4.8数据集的预处理

这里主要围绕邱日和周宽来展开，周宽协助邱日完成，具体处理方法我并不知道。

4.9 数据输入模型进行训练

把以上处理的数据加载入之前建立的模型中，就可进行模型的训练了，之后就可以得到我们自己模型的H5文件了。（由组长完成）

5.给我的分工

(1)学习python,了解python的基本语法，学会如何使用python;

(2)学习keras和tensorflow库的理论知识，知道模型构建的基本原理及神经网络的构建原理，自己学会跑keras的代码例子；

(3)有了以上的基础知识之后，学习预训练模型的基本知识，查看相关论文，学会预训练模型的加载原理，协助邱日组长加载预训练模型；

(4)了解迁移学习的基本概念以及过程，学会如何在预训练模型上进行迁移学习，构建神经网络，协助邱日组长在预训练模型上进行迁移学习；

(5)最后撰写SRT总研究报告和制作演讲PPT