### 关于食品数据库及python的云端实现调研报告

**目录：**

一：需求分析 ---------------------------------------------------------------------------------------------1

二：小组分工 ---------------------------------------------------------------------------------------------1

三：调研结果 ---------------------------------------------------------------------------------------------1

1：食品数据库 --------------------------------------------------------------------------------------1

2：Python的云端实现 -------------------------------------------------------------------------4

3：常见食品热量大小 -------------------------------------------------------------------------5

4：深度学习模型规划 -------------------------------------------------------------------------7

四：结果总结 ---------------------------------------------------------------------------------------------8

五：参考文献 ---------------------------------------------------------------------------------------------8

一 需求分析

我们的主要工作是通过云端传入的食品图像及重量，通过图像识别得到食品的种类，并根据其计算并返回食品的热量含量。

二 小组分工

一：调研阶段分工

食品数据库的调研：范博坤 王健军

Python的云端实现调研：高峰

常见食品热量调研：贾亚伟 王健军

二：实现阶段分工

图像预处理及计算阶段程序：范博坤

食品热量信息收集：贾亚伟

采用深度学习实现功能：王健军 范博坤

调用API实现功能：高峰 贾亚伟

三 调研结果

1. **食品数据库调研情况**

**（1）数据集寻找及对比过程**

最开始我们找到的数据集是Food-101 Dataset。该数据集由斯坦福大学于 2014 年发布，相关论文有《Food-101 — Mining Discriminative Components with Random Forests》。

Food-101 Dataset 是包含 101 中食品类别的图像数据集，主要用于图像分类，共有 101,000 张图像，每个类别的测试图像和训练图像分别有 250 张和 750 张，其中训练图像未经数据清洗，但所有图像均以重新进行尺寸缩放，最大边长在 512 像素。

但是我们下载了该数据集而后分析发现，该数据集主要包含的为西方的常见食品，且仅包含101种常见的食品。对于我们这次的需求来说并不太相符。而后又找到了一些关于印度的常见食物的数据集，同样不符合我们的要求。

最后我们找到了ChineseFoodNet。该数据集由美的全球人工智能团队提供，其中包括了共计208类中国各大常见菜系的实物和菜谱图片，合计180,000张，共18GB，是目前世界上最大的中国成品菜数据集。在下载评估过后，我们发现该数据集非常适合我们目前的情况。同时，通过对该数据集的研究我们发现关于食品热量的调研变得更加复杂，因为中国的成品菜成分并不单一，其单位质量的热量含量需要进行更细致的调研，这部分我们将在之后的工作中加强。

下表为ChineseFoodNet数据集的部分标签页：

|  |  |  |
| --- | --- | --- |
| 0 | 麻婆豆腐 | Mapo Tofu |
| 1 | 家常豆腐 | Home style sauteed Tofu |
| 2 | 煎豆腐 | Fried Tofu |
| 3 | 豆腐花 | Bean curd |
| 4 | 臭豆腐 | Stinky tofu |
| 5 | 酸辣土豆丝 | Potato silk |
| 6 | 土豆泥 | Pan fried potato |
| 7 | 香煎土豆 | Pan fried potato |
| 8 | 土豆焖豆角 | Braised beans with potato |
| 9 | 地三鲜 | Fried Potato, Green Pepper & Eggplant |
| 10 | 薯条 | French fries |
| 11 | 鱼香茄子 | Yu-Shiang Eggplant |
| 12 | 蒜泥茄子 | Mashed garlic eggplant |
| 13 | 肉末茄子 | Eggplant with mince pork |
| 14 | 辣白菜 | Spicy cabbage |
| 15 | 醋溜白菜 | Sour cabbage |
| 16 | 上汤娃娃菜 | Steamed Baby Cabbage |
| 17 | 手撕包菜 | Shredded cabbage |
| 18 | 蚝油生菜 | Sauteed Lettuce in Oyster Sauce |
| 19 | 炒青菜 | Saute vegetable |
| 20 | 炒空心菜 | tumis kangkung |

**（2）选择原因**

综合考虑我们选择了ChineseFoodNet数据集，一：数据量更多，食物的种类更多；二：更符合我们日常饮食情况，我们的饮食大部分都是中餐。但是选择了该数据集，意味着之后我们的对应食物的单位热量含量情况变得更加复杂。

**（3）数据集来源**

数据集来源：1><https://sites.google.com/view/chinesefoodnet/>

2>https://pan.baidu.com/s/1EgaDWW0fJVmBQ\_JWO64\_4Q（提取码4ma7）

**（4）ChineseFoodNet数据集使用情况**

快手科技曾经对该数据集进行过研究，采用了不同架构的神经网络来对该数据集进行训练和测试，并在不同深度上进行数据融合来找到最优的神经网络模型。[1]

实验结果如下表所示：

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Top\_1 Accuracy | Top\_1 Accuracy | Top\_5 Accuracy | Top\_5 Accuracy |
| Fusion Method |  |  |  |  |
| ResNet (18 + 34 + 50 + 152) | 79.19% | 79.46% | 96.03% | 96.16% |
| DenseNet (121 + 169 + 201) | 80.47% | 80.17% | 96.26% | 96.30% |
| ResNet (18 + 34 + 50 + 152) + Densenet (121 + 169 + 201) | 80.89% | 81.08% | 96.60% | 96.67% |
| ResNet (18 + 34 + 50 + 152) + Densenet (121 + 169 + 201) + VGG19-BN | 81.23% | 81.12% | 96.79% | 96.76% |
| ResNet152 + DenseNet (121 + 169 + 201) + VGG19-BN | 81.43% | 81.55% | 96.73% | 96.76% |

并由此得出以下结论，一，通过在不同深度上应用数据融合方法比使用单一深度网络拥有更高的整体性能；二，不同网络体系结构的组合比相同网络体系结构的组合在提高性能方面表现出更大的优势，深度和广度网络的组合获得了最好的结果。

浙江大学的苏国炀也利用该数据集进行了训练及测试，提出了一种基于双线性结构的双流卷积神经网络,用以识别单目标中餐菜品图像。通过具有不同视野的双流设计提取置信度更高的中餐菜品图像特征表示,通过双线性结构实现纹理特征的平移不变性,利用紧凑的双线性池化降低网络模型的复杂度,利用加法跨流连接方式实现双流特征信息的融合。针对训练过程中数据集各类别图片数量不均衡的问题,提出权重交叉熵函数,并通过计算混淆损失减少网络过拟合现象。在包含208类中餐菜品的公开数据集ChineseFoodNet上的测试结果表明,本文所提网络的Top-I准确率为81.16%,Top-5准确率为96.32%,显著优于使用现有经典网络的识别准确率。

通过该数据集训练的模型及其他的一些目标提取和分割的功能，他实现了对食堂中餐菜品的分割和识别。[2]

该数据集是由美的团队提供的，是美的在往智能家居方向转型时所发布的，但是我在查找了美的智能家居的产品却没有发现与该数据集相关的，最为相似的是通过对食材图像的识别来提供菜谱。

1. **Python的云端实现调研情况**
2. 安装前的准备

1.配置服务器程序

现在，我们用ubuntu切换到的root用户去进行接下来的操作。

首先是配置Supervisor，我们进入到Supervisor的配置目录：

cd /etc/supervisor/conf.d/

同样，我们创建一个配置文件jian.conf：

vim jian.conf

并输入以下内容：

[program:jian]  
command=/data/web/ZhihuJianMingJiaoCheng/env/bin/gunicorn run:app -c/data/web/ZhihuJianMingJiaoCheng/gunicorn.conf  
directory=/data/web/ZhihuJianMingJiaoCheng  
user=www  
autostart=true  
autorestart=true  
stdout\_logfile=/data/web/ZhihuJianMingJiaoCheng/logs/gunicorn\_supervisor.log  
environment = MODE="PRODUCTION"

同样，创建一个配置文件jian.conf:

vim jian.conf

并输入以下内容：

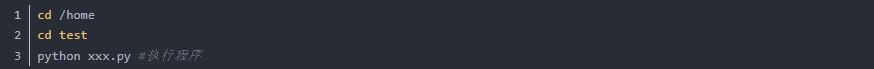
server {  
listen 80;  
server\_name [www.xxx.com](https://link.zhihu.com/?target=http%3A//www.xxx.com/" \t "_blank);  
  
root /data/web/ZhihuJianMingJiaoCheng/;  
access\_log/data/web/ZhihuJianMingJiaoCheng/logs/access.log;  
error\_log /data/web/ZhihuJianMingJiaoCheng/logs/access.log;  
  
location / {  
proxy\_set\_header X-Forward-For $proxy\_add\_x\_forwarded\_for;  
proxy\_set\_header Host $http\_host;  
proxy\_redirect off;  
if (!-f $request\_filename) {  
proxy\_pass [http://unix](https://link.zhihu.com/?target=http%3A//unix/" \t "_blank):///tmp/jian.sock;  
break;  
}

二、配置python环境

首先配置python3.6的环境，执行以下linux命令配置python3.6环境：



1. **把python3项目上传到linux的云服务器的/home中**
2. **执行python程序**



4.返回值返回到后端

1. **常见食品热量情况**

关于常见食品的热量，我们前期收集各大网站与APP中关于各种食物热量的数据，通过筛选找出了一家食物热量种类比较齐全的做关于饮食健康的网站，使用网站爬虫技术来将各种食物的热量数据进行收集汇总，一共收集了542种常见食材的热量含量。

部分数据如下:

|  |  |
| --- | --- |
| 食物名称 | 热量/100g |
| 鸡蛋 | 热量：139 大卡(每100克) |
| 猪小排 | 热量：295 大卡(每100克) |
| 鸡(土鸡，家养) | 热量：124 大卡(每100克) |
| 鸡(母鸡，一年内) | 热量：256 大卡(每100克) |
| 鸡(肉鸡，肥) | 热量：389 大卡(每100克) |
| 虾 | 热量：93 大卡(每100克) |
| 鸭肉 | 热量：240 大卡(每100克) |
| 猪蹄 | 热量：260 大卡(每100克) |
| 猪肉(瘦) | 热量：143 大卡(每100克) |
| 火腿肠 | 热量：212 大卡(每100克) |

但是为了和我们数据集的情况匹配，我们需要通过菜谱等资料，来合成计算对应菜品的热量含量。

下表展示的是我们收集的部分数据：

|  |  |  |
| --- | --- | --- |
| 食物名称 | 食物英文名称 | 食物热量/100g |
| 麻婆豆腐 | Mapo Tofu | 122大卡 |
| 家常豆腐 | Home style sauteed Tofu | 120大卡 |
| 煎豆腐 | Fried Tofu | 92大卡 |
| 豆腐花 | Bean curd | 15大卡 |
| 臭豆腐 | Stinky tofu | 181大卡 |
| 酸辣土豆丝 | Potato silk | 101大卡 |
| 土豆泥 | Pan fried potato | 124大卡 |
| 香煎土豆 | Pan fried potato | 129大卡 |
| 土豆焖豆角 | Braised beans with potato | 146大卡 |
| 地三鲜 | Fried Potato, Green Pepper & Eggplant | 64大卡 |
| 薯条 | French fries | 226大卡 |
| 鱼香茄子 | Yu-Shiang Eggplant | 78大卡 |
| 蒜泥茄子 | Mashed garlic eggplant | 51大卡 |
| 肉末茄子 | Eggplant with mince pork | 219大卡 |
| 辣白菜 | Spicy cabbage | 61大卡 |
| 醋溜白菜 | Sour cabbage | 46大卡 |
| 上汤娃娃菜 | Steamed Baby Cabbage | 36大卡 |
| 手撕包菜 | Shredded cabbage | 56大卡 |
| 蚝油生菜 | Sauteed Lettuce in Oyster Sauce | 31大卡 |
| 炒青菜 | Saute vegetable | 37大卡 |
| 炒空心菜 | tumis kangkung | 41大卡 |

**4>食品数据库调研情况**

为了更好的完成识别任务，得到一个有效的实验结果，我们将尝试传统的卷积神经网络模型以及新颖的Transformer网络结构。

对于传统的卷积神经网络，我们以ImageNet比赛中各种经典的网络模型作为参考，力求做到近年来CVPR中所能达到的结果。并不断的优化模型，不但要求网络能够识别的准，更能快速的实现识别。力争做到有准又快。预计我们将采用Resnet[3]，Resnext[4]，mobilenetV3[5]，Seresnet[6]这四种网络结构分别做对比实验，以便得到最优的卷积模型结构。

对于近年来及其风靡且有效的Transformer网络结构，我们重点复现CVPR2021中，华为诺亚实验室提出的同时利用了块内部序列和块之间序列信息的transformer模型，称之为Transformer-in-Transformer[7]（简称TNT）。以及CVPR 2021中由微软亚洲研究院提出的一种称为Swin Transformer[8]的新型视觉Transformer。该模型结构已经取得了目标检测，实例分割，语义分割三项比赛第一。

四 结果总结

目前我们选定了ChineseFoodNet数据集，该训练集每种食物有300到1500张不等的图片。我们之后将参考其他已实现的项目，对其进行复现，初步实现菜品图片的识别分类功能。

与此同时，我们还将寻找是否存在成熟的可调用的API，两个方向同步进行。并继续收集完善菜品的热量数据，最后通过算法计算出食物的总热量，并将结果返回到云端。

五 参考文献

1. Xin Chen,Yu Zhu,Hua Zhou.ChineseFoodNet: A Large-scale Image Dataset forChinese Food Recognition.https://www.researchgate.net/publication/316779861\_ChineseFoodNet\_A\_large-scale\_Image\_Dataset\_for\_Chinese\_Food\_Recognition
2. 苏国炀. 基于图像的中餐菜品分割与识别[D].浙江大学,2019.
3. Kaiming He,Xiangyu Zhang,Shaoqing Ren ,Jian Sun.Deep Residual Learning for Image Recognition.https://arxiv.org/pdf/1512.03385.pdf
4. Saining Xie,Ross Girshick,Piotr Dollar, Zhuowen Tu,Kaiming He.Aggregated Residual Transformations for Deep Neural Networks.https://arxiv.org/pdf/1611.05431.pdf
5. Andrew Howard, Mark Sandler, Grace Chu.Searching for MobileNetV3.https://arxiv.org/abs/1905.02244v2
6. Kaiming He,Xiangyu Zhang,Shaoqing Ren,Jian Sun.Deep Residual Learning for Image Recognition.https://arxiv.org/pdf/1512.03385.pdf
7. Kai Han,An Xiao,Enhua Wu.Transformer in Transformer.https://arxiv.org/pdf/2103.00112.pdf
8. Ze Liu, Yutong Lin, Yue Cao.Swin Transformer: Hierarchical Vision Transformer using Shifted Windows.https://arxiv.org/abs/2103.14030