文件编号：Calorier-SWC2020-20200184

受控状态：■受控 □非受控

保密级别：□公司级 □部门级 ■项目级 □普通级

采纳标准：CMMI DEV V1.2





卡路里

**Calorier**

技术研究报告

**Version 1.2.0**

2019.11.26

**Written by Calorier team**



**All Rights Reserved**

目录

[1 问题聚焦 1](#_Toc22848191)

[1.1 问题描述 1](#_Toc22848192)

[1.2 问题抽象 1](#_Toc22848193)

[1.3 问题定位 1](#_Toc22848194)

[1.4 问题评估 1](#_Toc22848195)

[1.5 问题分解 1](#_Toc22848196)

[2 相关工作 1](#_Toc22848197)

[3 技术方案 1](#_Toc22848198)

[3.1 技术方向 1](#_Toc22848199)

[3.2 技术选择 1](#_Toc22848200)

[3.3 结果期望 1](#_Toc22848201)

[4 技术实践 1](#_Toc22848202)

[4.1 使用的开发框架及依赖的Library 1](#_Toc22848203)

[4.2 技术实践过程 1](#_Toc22848204)

[5 结果验证 1](#_Toc22848205)

记录更改历史

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **序号** | **更改原因** | **版本** | **作者** | **更改日期** | **备 注** |
| 1 | 增加问题聚焦内容 | 1.0.0 | 张胤泽 | 2019/11/23 |  |
| 2 | 增加相关工作内容 | 1.1.0 | 张胤泽 | 2019/11/24 |  |
| 3 | 增加技术方案内容 | 1.2.0 | 张胤泽 | 2019/11/26 |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |

# 问题聚焦

## 问题描述

目前市面上可以做食物图像识别得到每种食物重量及营养素含量的算法或者软件很少，有的识别算法或者软件对食物重量的识别的准确度只能达到50％左右，所以在用户想通过食物图像识别获取食物重量及营养素含量的时候，误差是可能达到40％以上。食物重量的识别的准确度只能达到50％左右。

## 问题抽象

机器学习有两个大的方面，一个是做分类，一个是做回归。分类是一个基本问题。例如，当我们知道有不同颜色表示的两类物体，在向量空间中，我们有一种什么样的算法能够把这两类物体分开，并且分得更好。本问题旨在提高食物图像识别的精度，即如何提高机器学习分类算法的准确率。

## 问题定位

本问题属于机器学习图像识别领域。随着大数据深度模型以及硬件设备的完善,深度学习技术也得到了广泛发展,促进了计算机视觉、语音识别、自然语言处理等领域的发展,而这些模型的成功非常依赖于大量数据。本项目的研究基于food-101数据集通过机器学习的方法来进行图像识别。

## 问题评估

机器学习是一个比较基础的研究，模式识别、数据挖掘中的很多方法都是来自于机器学习，模式识别和数据挖掘可以用在语言、文字、图像和视频识别中。这里列的都是大家常见的一些应用，其实还包括生物、医学、材料科学等其他学科方面的应用。机器学习是一个基础的研究工具，或者说是理论和方法。所以在大数据时代，机器学习是核心方法之一。这样一个方法在任何一个大的应用领域都有可能有发挥它的用武之地。

深度学习在图像处理中的优势深度学习的优势

（1）从统计，计算的角度看，DL特别适合处理大数据这个大数据是除了数量上的大，还有更重要的是维度的大，很多算法本身是无法处理高纬度数据的，例如Kernel学习机相关的算法， 虽然理论上是先将数据向高维空间映射，然后在高维空间进行线性的求解，实际上在处理的时候还是回到原空间处理。传统的BP算法针对高维的数据也是效果不佳。

CNN等为什么对图像领域更加有效，因为其不但关注了全局特征，更是利用了图像识别领域非常重要的局部特征，应该是将局部特征抽取的算法融入到了神经网络中。图像本身的局部数据存在关联性，而这种局部关联性的特征是其他算法无法提取的。深度学习很重要的是对全局和局部特征的综合把握

（2）深度学习不是一个黑箱系统。它像概率模型一样，提供一套丰富的、基于联接主义的建模语言。利用这套语言系统，我们可以表达数据内在的丰富关系和结构。比如用卷积处理图像中的二维空间结构，用递归神经网络（Recurrent Neural Network）处理自然语言等数据中的时序结构

（3）深度学习几乎是唯一的端到端的学习系统，它直接作用于原始数据，自动逐层进行特征学习，整个过程直接优化目标函数。

随着深度学习模型的不断发展，深度学习在图像处理领域的应用也越来越多，这个方面研究的热度也很高。

深度学习在图像识别中的发展趋势

（1）模型层次不断加深

（2）模型结构日趋复杂，传统的卷积神经网络都是简单的 conv-pool-FC，后来NIN 用mlpconv 代替传统的 conv层（mlp 实际上是卷积加传统的多层感知器）。这样做一方面降低过拟合程度提高模型的推广能力，另一方面为大规模并行训练提供非常有利的条件

（3）海量的标注数据和适当的数据扰动DL需要大量的数据，现有的图像数据不能满足需求，结合图像数据的特点，通过平移、水平翻转、旋转、缩放等数据扰动方式可以产生更多的有效数据，普遍提高识别模型的推广能力。

## 问题分解

图像识别技术的过程分以下几步：信息的获取、预处理、特征抽取和选择、分类器设计和分类决策。信息的获取是指通过传感器，将光或声音等信息转化为电信息。也就是获取研究对象的基本信息并通过某种方法将其转变为机器能够认识的信息。预处理主要是指图像处理中的去噪、平滑、变换等的操作，从而加强图像的重要特征。

特征抽取和选择是指在模式识别中，需要进行特征的抽取和选择。简单的理解就是我们所研究的图像是各式各样的，如果要利用某种方法将它们区分开，就要通过这些图像所具有的本身特征来识别，而获取这些特征的过程就是特征抽取。在特征抽取中所得到的特征也许对此次识别并不都是有用的，这个时候就要提取有用的特征，这就是特征的选择。特征抽取和选择在图像识别过程中是非常关键的技术之一，所以对这一步的理解是图像识别的重点。

分类器设计是指通过训练而得到一种识别规则，通过此识别规则可以得到一种特征分类，使图像识别技术能够得到高识别率。分类决策是指在特征空间中对被识别对象进行分类，从而更好地识别所研究的对象具体属于哪类。

# 相关工作

技术方案一：

由Nidhin Pattaniyil 和Reshama Shaikh在2016年完成并持续维护。

他们的解决方案是：

用fastai库训练一个给食物照片分类的深度学习模型

用Heroku和Flask将这个模型部署到网页和移动端

基本的实现步骤分为三部分：

第一部分：训练分类器

第二部分：将模型部署到网页端

第三部分：将模型部署到移动端

Github地址:

https://reshamas.github.io/deploying-deep-learning-models-on-web-and-mobile

技术方案二：

于2017年03月22日实现

解决方案：

在对一个预先训练过的Google InceptionV3模型进行微调之后来训练food-101模型。

结果：

可在测试集上使用每项一次裁剪就可以达到82.03%的Top-1精度。以10种作物为例，以最常见的预测类别为例，可以达到86.97%的Top-1精度和97.42%的Top-5精度。

Github地址:

<https://github.com/stratospark/food-101-keras>

技术方案三：

2017年麻省理工，图片转食谱。

实现方案：

Recipe1M+，这是一个由超过100万个烹饪食谱和1300万个食物图像组成的大型结构化语料库。作为配方数据的最大公共可用集合，Recipe1M+提供了在对齐的多模式数据上训练高容量模型的能力。利用这些数据，训练了一个神经网络来学习如何将配方和图像联合嵌入到图像配方检索任务中，从而得到压缩结果。此外，论文还证明，通过增加一个高层次分类目标的正则化，既可以提高检索性能，与人类的检索性能相媲美，又可以实现语义向量算法。

链接地址：<http://pic2recipe.csail.mit.edu/>

# 技术方案

## 技术方向

我们选择的技术方向为深度学习。深度学习与传统模式识别方法的最大不同在于它所采用的特征是从大数据中自动学习得到，而非采用手工设计。好的特征可以提高模式识别系统的性能。过去几十年，在模式识别的各种应用中，手工设计的特征一直处于统治地位。手工设计主要依靠设计者的先验知识，很难利用大数据的优势。由于依赖手工调参数，因此特征的设计中所允许出现的参数数量十分有限。深度学习可以从大数据中自动学习特征的表示，可以包含成千上万的参数。

## 技术选择

我使用python语言调用了pytorch机器学习框架的MobileNet-v2模型实现了食物的识别分类。我对MobileNet-v2的训练参数进行了修改以使模型有更高的识别率，因为本次使用的训练集为food-101训练集，与MobileNet-v2的输入参数不同，所以我首先对图片数据进行了预处理。

## 结果期望

食物识别率达到85%以上。

# 技术实践

## 使用的开发框架及依赖的Library

## 技术实践过程

# 结果验证