EatRight: An Instance Segmentation based Calorie Counter

Paavni Shukla, Jagriti Dixit

First Year, MS, Electrical Engineering

Introduction

The statistics on world obesity rates indicate that the worldwide prevalence of obesity nearly tripled between 1975 and 2016. As a result of this, a lot of people constantly track their calorie intake throughout the day. Our project aims to simplify this task to just clicking a picture for the users to find the number of calories present in the food being consumed.

While a lot of off the shelf calorie counters exist today most of them rely on manual input of the food item and its quantity. Our project aims to do Instance segmentation based on the food image and identify its category and quantity. This would enable the user to upload a photo and identify the food item and its calories.

Fig 1: Food Images and their corresponding masks

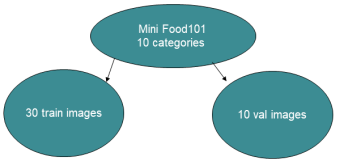
Problem Statement

The input to our algorithm is an image of the food. We then use Mask R-CNN deep learning model to output a predicted calorific value of the food item. This will be done by first identifying the Region of Interest(ROI), wherein the food items are present followed by its instance segmentation telling us the labels of all food items in the image. This is followed by the calorie calculation based on the number of food items of each type present in the image and the total calories is the image is predicted.

To track the quality of our predictions, we are using mean Average Precision for all images at an IoU of 50%. We are also tracking the train and the validation accuracy and their trends to minimize validation loss and overfitting.

Dataset

Mini Food101- custom made dataset with manual annotation for 10 categories of Food101 with 40 images each



Methods

- Baseline Model- Classification of images based on food categories, Resnet-50
- Instance Segmentation- Mask RCNN with transfer learning from MS COCO + Calorie Lookup based on number of instances of each category

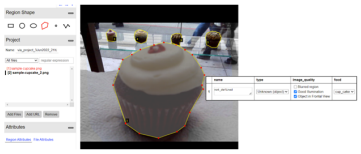


Fig 2:Manual annotation using VGG tool

Fig 3:Left to right, top to bottom- Top anchors before refinement, top anchors after refinement and NMS, Final proposal by RPN, Final step of object detection

Experiments & Analysis

Resnet50 – Baseline Model

Class	Test Accuracy
caesar_salad	96.80%
cup_cakes	97.20%
donuts	92.00%
hot_dog	96.80%
macaroni_and_cheese	97.20%
macarons	96.80%
omelette	94.00%
pizza	98.00%
prime_rib	97.20%
samosa	96.80%

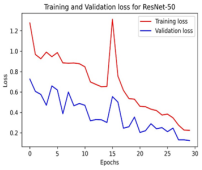


Fig 4: Category-wise accuracy with Resnet50, train vs val loss curve

Mask RCNN + Calorie Lookup

Fig 6: train vs val loss Mask RCNN

Metric	Value
Mean Average Precision (mAP@IOU=50)	0.807
Mean Calorific Accuracy	76.03%

Fig 5:mAP and mCA metrics

Fig 7:Confusion matrix for cupcake image

Conclusions & Future Work

Mask RCNN can be utilized towards better and accurate calorie lookups.

- Use scripts for automating the manual annotation task
- Model can be scaled to include more food categories
- Changing the backbone from Resnet101 to DenseNet101 or Inception networks as they have recorded highest performances for Food101
- Using pretrained weights from the baseline model of Resnet-50
- End to End UI so that real time image uploading and calorie lookup can be done
- Volumetric analysis for food items which can't be segregated based on instances like Mac and Cheese , Caesar salad