



HACETTEPE UNIVERSITY

DEPARTMENT OF COMPUTER  
ENGINEERING  
SPRING 2019

---

## BBM 416 Computer Vision Project Progress Report

---

*Supervisors:*

Dr. Nazlı İkizler Cinbis

TA Özge Yalcinkaya

*Names:*

Abdülşemet Kalkan

Ege Berke Balseven

Yiğit Burdurlu

*Subject:*

Food Recognition for Calorie

Management

*Due Date:*

19.05.2019

# Food Recognition for Calorie Management

Abdölsamet Kalkan

21427033

Ege Berke Balseven

21590776

Yiğit Burdurlu

21426735

b21427033@cs.hacettepe.edu.tr

Hacettepe University

b21590776@cs.hacettepe.edu.tr

Ankara, TURKEY

b21426735@cs.hacettepe.edu.tr

May 19th, 2019

## Abstract

Our paper address the issue of how much calorie food is. For this purpose, we propose a CNN-based food classification and food detection model using Pytorch to estimate the number of calories in an image of food taken; based on the average calorie value of each detected category. Our dataset contains 101.000 images of 101 different types of food.

## 1 Introduction

Obesity in the world is increasingly rising. One of the biggest reasons is an unawareness of calorie of those what people eat. People do not know calories of their meal so that they eat unconsciously and gain weight readily. We aim to address this issue by means of Deep Learning. Our solution will recognize food in a meal and calculate total calories.

There are a few big challenges to solve this problem. First one is to identify food or ingredients in it. Identifying ingredients would give more precise results to calculate calorie. But it is immensely hard to identify due to the smallness of those. Hence, we have ended up with classifying food instead. However, a new problem comes out. Although food recognizing is easier, calculating the calorie is hard. Because it is not known the portion of food on the plate. But averagely gives more precise results. That is why we have chosen food based classification.

## 2 Related Works:

- Pouladzadeh et al.(4) proposed an approach that segments the image and predicts the calorie of the food in the given image. They used K-Mean clustering for the image segmentation and find the food portions in the image. For classification Support Vector Machines(SVM) is used. The SVM model was trained in order to classify the food that given. SVM model took Color Features(10 categories), Size Features(6 categories), Shape Features(5 categories), Texture Features(5 categories) and classify the food. They needed two images(one is from the side and other is from the top) which are contains food and the user's thumb for the volume estimation. They reached 92.21% of accuracy.
- Chokr et al.(2) used a dataset of fast-food images based on the Pittsburgh Fast-Food Image Dataset. Mathworks Image Processing Toolbox is used for extracting features. Total of 11,868 raw features extracted from RGB representation of the image. Raw features reduced using Principal Component Analysis(PCA) and Information Gain(InfoGain) to 23 features. Using these features and Sequential Minimal Optimization(SMO) they classified the food. Size prediction is done by using Random Forest in grams. Finally, the calorie of the food is predicted using Multilayer Perceptron. They experiment on different types of representation of an image such as Averaged RGB, Gray Scale, BW 0.7 and BW 0.5 but they get the best results with RGB representation.
- Liang et al. (3) used a dataset of 2978 images of 19 different food with a single yuan coin. They used Faster CNN for object detection. Each bounding box that created by Faster CNN is classified. They used the GrabCut algorithm for image segmentation. They divided food into 3 categories ellipsoid, column and irregular. Using the coin as a reference point they calculate the volume of the food depending on its shape. Knowing the volume mass can be easily calculated using the density of the food. With the mass is known they calculated the calorie of the food.

## 3 Method:

- During pre-process of project, we had to design a solution for the classification problem. After a brief search, we came up with a solution which uses neural networks because neural networks give outputs which refer to scores of all the classes.Also, in state-of-art they are pretty successful classifiers, and are broadly used. Moreover, with convolutional neural network layers, we finalized

our solution in a perfect fit. Thus, we decided the architecture of our neural network as sequentially contains convolutional and fully connected layers.

- We end up the number of convolutional layers as 3 because in the progress period, we wanted to have some feedback and results from what we had designed so we chose a number which provides as quickly as possible and also successful training phase for the layers of neural network. Thus, the number of fully connected layers was also decided as 1. The architecture of the our neural network is shown in figure 1.
- Convolutional layers provide a role which extracts features of images by the help of kernels. These features enter into the fully connected layers as inputs.
- In this progress period, we have used cross entropy loss and Adam optimizer for now, however until the final report, we set our minds to implement various optimizer and loss functions as well.
- Below, we show train loss and accuracy table. In final, we are going to add confusion matrix of classes, validation loss and accuracy, Top-5 accuracy, average calories for each 101 foods.

## 4 Initial Experimental Settings:

We used data set of 101 food categories, with 101.000 images. (1).For each class, 204 manually reviewed test images are provided as well as 796 training images. Those images are from "foodspotting.com" which is one of the well-known sites in which people share the images of what they eat and give information about.

our parameters:

Epochs: 15

Classes: 101

Batch size: 64

Learning rate: 0.001

Also we resized images to 128x128.

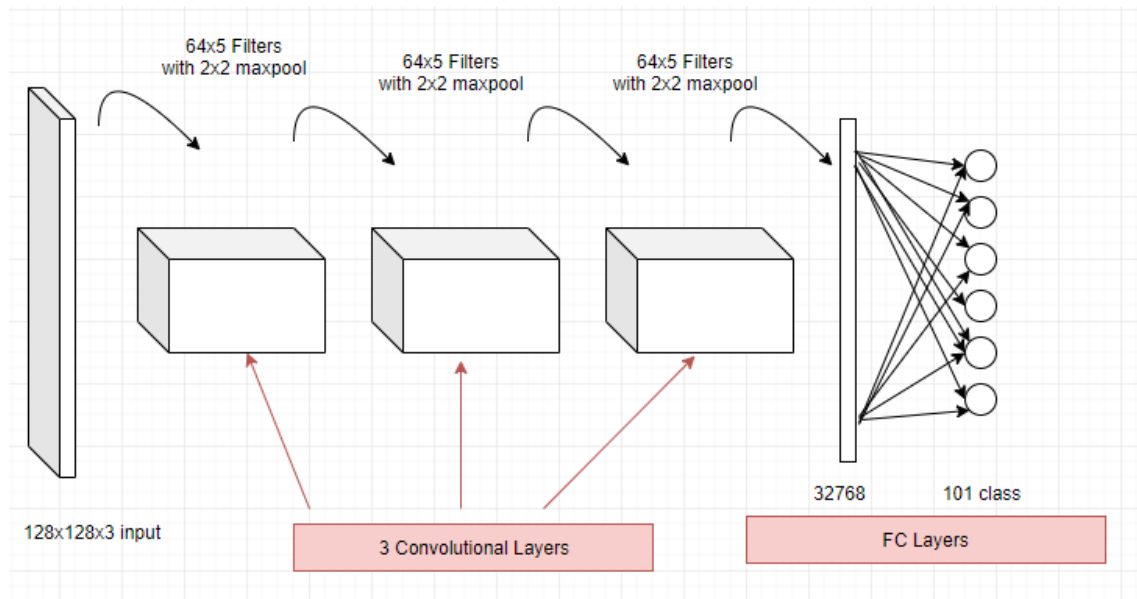
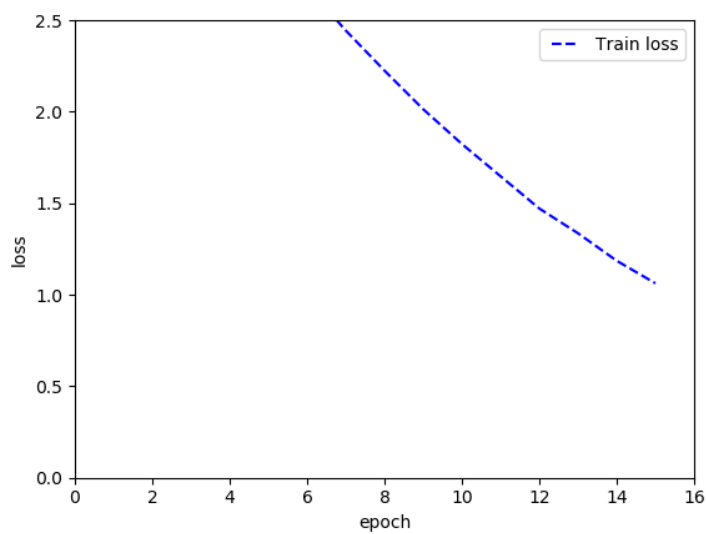


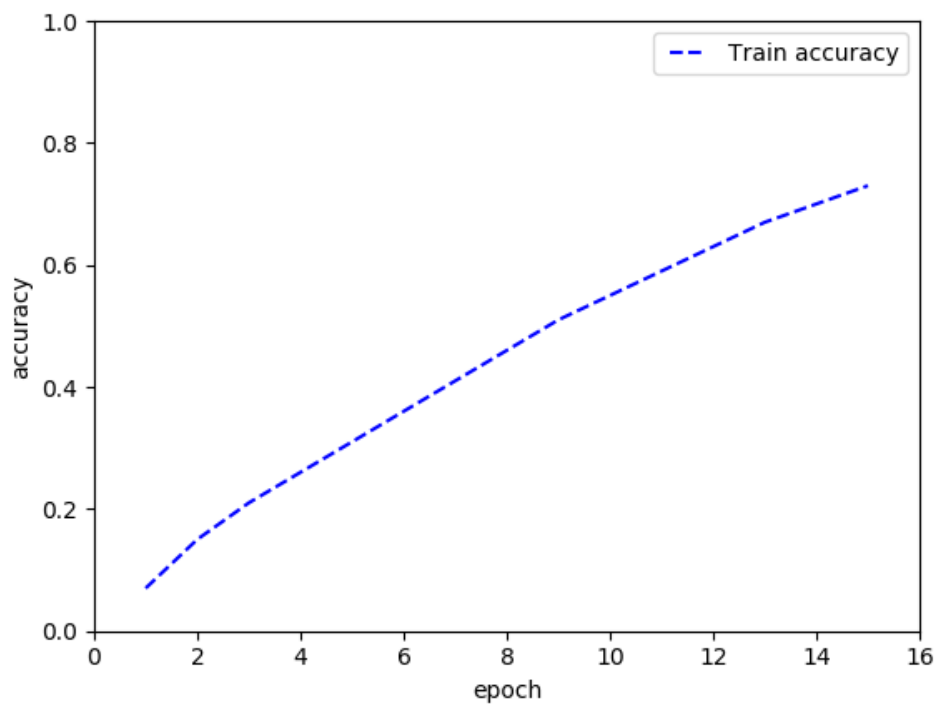
Figure 1: Our Neural Network

## 5 Initial results:

Train loss:



**Train accuracy:**



As it is seen from accuracy and loss graphs, results are getting better according to increase of epoch number.

**Test accuracy:**

```
20 # Hyper parameters
21 num_epochs = 15
22 num_classes = 101
23 batch_size = 64
24 learning_rate = 0.001
```

```
Run: Main x
Epoch 15, Batch 1472, Loss: 1.2069740
Epoch 15, Batch 1536, Loss: 1.2196032

Train set = Epoch: 15   Accuracy 74366/101056 (73%) Average loss: 1.0623274

Overall test Accuracy : 79.83886623956514 %
Program is finished in 16150.18192243576 seconds.
```

Overall test accuracy is not bad. But bigger epoch numbers and convolutional, fc layers are gets us better results.

## References

- [1] L. Bossard, M. Guillaumin, and L. Van Gool. Food-101 – mining discriminative components with random forests. In *European Conference on Computer Vision*, 2014. [2](#)  
[3](#)
- [2] S. E. M. Chokr. Calories prediction from food images. [2](#)
- [3] J. L. Y. Liang. Computer vision-based food calorie estimation: Dataset, method, and experiment. [2](#)
- [4] S. S. P. Pouladzadeh and R. Al-Maghrabi. Measuring calorie and nutrition from food image. In *Instrumentation and Measurement, IEEE Transactions on*, 63:1947– 1956, 08 2014. [2](#)