# CSCI - GA 3033 - 91

# Introduction to Deep Learning Systems

# Final Project Report

Hanwei Peng (hp2166)

Zijing Sun (zs2198)

May 14, 2021

## Project Description

Our project topic is to train a food image classifier and to implement a food nutrition analyzer. The user can upload the photo of food and our system will recognize what the food is and analyze its nutrition details, such as its Energy, Vitamins, Minerals and its nutrition score. To train the food image classifier, we used the Food-101 dataset which has one hundred and one types of food and we trained a ResNet50 model and a Inception V3 model to compare their accuracy and choose the one with the highest performance. To implement the food nutrition analyzer, we used the Open Food Facts dataset and recreated an excel file to take out the nutrition information we want and correspond them with food names. Finally, we will output a visualization of the nutrition details for a given food picture, including a pie chart showing the vitamin percentage and a pie chart showing the micronutrients, as well as a nutrition score and calorie.

## Approach

- Collect data

  From Open Food Facts and HappyForks.com these two organizations, we collected the Energy, Vitamins, Minerals facts of each type of food we need and corresponded them with the name of food. We organized these data into an Excel file for later search. Here is a screenshot of the nutrition detail table we collected.

| name | Energy kcal/100 g | Vitamin C | Vitamimn B | Vitamimn B2 | Vitamimn B3 | Vitamimn B5 | Vitamimn B6 | Vitamin B12 | Vitamin A | Vitamimn E | Vitamimn D | Vitamimn K | Calcium mg | Magnesium | Phosphorus | Iron | Potassium |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| apple_pie | 258 | 866 | | 27 | 263 | 119 | 38 | | 3.37 | 1.52 | | | 10 | 0.182 | 24 | 0.776 | 65 |
| baby_back_ribs | 229 | | | 309 | 500 | | 487 | | | | | | 29 | 22 | 163 | 1.29 | 225 |
| baklava | 485 | 130 | | 203 | 1895 | 230 | 111 | 1 | 5 | 2280 | | | 40 | 40 | 115 | 6.6 | 4.6 |
| bread_pudding | 111 | 1000 | 33 | 171 | 136 | 310 | 32 | 0.37 | | 30 | | 0.2 | 112 | 21 | 94 | 0.34 | 156 |
| breakfast_burrito | 221 | 3000 | 287 | 143 | 248 | 403 | 85 | 0.08 | | 37 | | 7.6 | 52 | 31 | 117 | 2.51 | 210 |
| bruschetta | 258 | | | | | | | | | | | | | | | | |
| caesar_salad | 71 | 3500 | 89 | 107 | 504 | 201 | 76 | 0.11 | 7.6 | 150 | | 89.1 | 140 | 18 | 95 | 1.04 | 230 |
| caprese_salad | 141 | 8600 | 51 | 105 | 411 | 174 | 74 | 0.64 | 0.88 | 910 | | 27.2 | 213 | 16 | 161 | 0.32 | 176 |
| carrot_cake | 415 | 1200 | 265 | 170 | 2205 | 294 | 79 | 0.04 | 1.93 | | | | 172 | 8 | 247 | 1.8 | 169 |
| cheesecake | 321 | 400 | 28 | 193 | 195 | 517 | 52 | | 0.54 | 560 | | 4.4 | 51 | 11 | 93 | 0.63 | 90 |
| chicken_curry | 80 | 8900 | 58 | 82 | 1477 | 200 | | | | 780 | | | 30 | 20 | 61 | 0.72 | 214 |
| chicken_quesadilla | 312 | 7200 | | 226 | 191 | 2408 | 97 | 0.24 | | 490 | | 6.8 | 268 | 22 | 269 | 2.17 | 151 |
| chicken_wings | 211 | | 74 | 242 | 5500 | 545 | 142 | 0.4 | 0.058 | 360 | | | 28 | 20 | 205 | 2.41 | 190 |
| chocolate_cake | 389 | 100 | | 57 | 773 | 190 | | | | 433 | | 28.9 | 30 | 31 | 137 | 3.04 | 270 |
| chocolate_mousse | 260 | 500 | 51 | 147 | 595 | 202 | 29 | 0.21 | 0.413 | | | | 77 | 32 | 231 | 1.08 | 285 |
| churros | 444 | | 163 | 125 | 1421 | | 11 | | | 1650 | | 13.9 | 7 | 6 | 30 | 1.27 | 30 |
| clam_chowder | 61 | 2100 | 82 | 173 | 813 | 297 | 70 | 4.93 | | | | | 68 | 12 | 175 | 1.24 | 181 |
| crab_cakes | 415 | 1200 | 265 | 170 | 2205 | 294 | 79 | 0.04 | 1.93 | | | | 172 | 8 | 247 | 1.8 | 169 |
| creme_brulee | 257 | 500 | | | | | | | 0.5 | | | | 150 | | | | |
| deviled_eggs | 198 | | 61 | 468 | 58 | 1286 | 111 | 1.02 | 0.479 | 1210 | 0.08 | 13.9 | 46 | 9 | 158 | 1.1 | 116 |
| donuts | 426 | 100 | 233 | 198 | 1512 | 435 | 27 | 0.24 | | | | | 60 | 17 | 117 | 1.06 | 102 |
| dumplings | 152 | 1500 | 207 | 156 | 1793 | 256 | 71 | 0.04 | | | 0.12 | 2.4 | 39 | 12 | 63 | 1.5 | 98 |
| edamame | 122 | 6100 | 200 | 155 | 915 | 395 | 100 | | | 680 | | 26.7 | 63 | 64 | 169 | 2.27 | 436 |
| falafel | 235 | 1500 | 139 | 145 | 944 | | 122 | | | 2110 | | 24.5 | 55 | 77 | 189 | 2.94 | 402 |
| french_fries | 260 | 6300 | 74 | 31 | 2241 | 540 | 180 | | | 100 | | | 9 | 18 | 77 | 1.32 | 430 |

- Manage Input data

  We downloaded the Food101 dataset and divided the dataset to train and test folder. Reduced the input dataset size to a quantity so that the training time won't last for too long. We decreased the input classes from 101 to 60 and limited the food image size to 299x299.

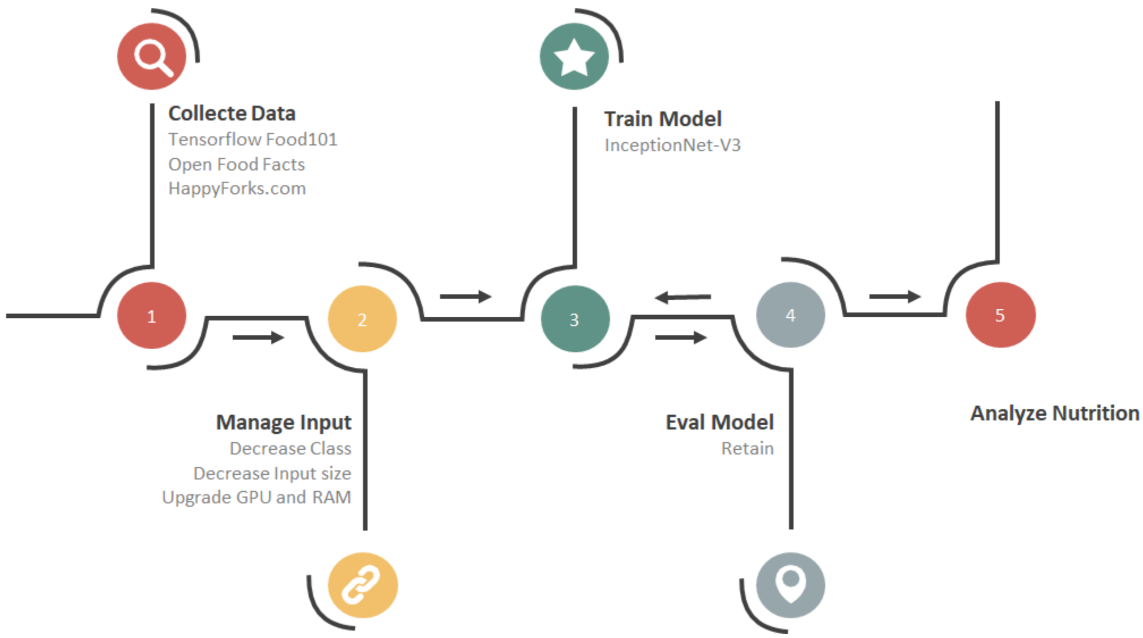
- Train and Eval Model

  We wanted to train a model that having an food image as input, will output the name of food with high accuracy. Because Food101 is a larger dataset whose size is 5 GB, we used services provided by Google Colab Pro (RAM 25GB, V100 GPU) and Google Cloud Platform (RAM 29GB, V100 GPU) to train our model. During the training, we adjusted hyperparameters like learning rate, epoch and used SGD optimizer and Dropout to improve the accuracy. Also, we trained both Resnet50 and Inception V3 models, because after some research [5], we were concerned that Resnet50 accuracy may can't reach our expectations. Finally, we compared the accuracy from the Resnet50 model and Inception V3 model and chose the model with the best accuracy as our final model.

- Analyze Nutrition

  After getting the predicted name of food from our final model, we used the Excel file we created before to retrieve the nutrition details according to the food name. And we used pie charts to visualize the percentage of vitamins and minerals.

## Solution Diagram




## Implementation Details

Dataset: Tensorflow Food101, Open Food Facts, HappyForks.com

FrameWork: Pytorch 1.8.1, Scipy 1.1.0, Keras 2.4.3, TensorFlow 2.4

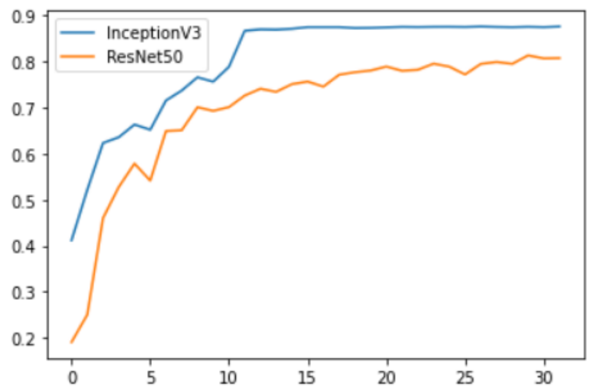
Model: ResNet50, Inception V3

Platform: Colab Pro (RAM 25BG, GPU V100), GCP (RAM 26BG, GPU V100)

When Colab Pro reached GPU limitation, we changed our platform to Google Cloud Platform to continue training.

## Experimental Results

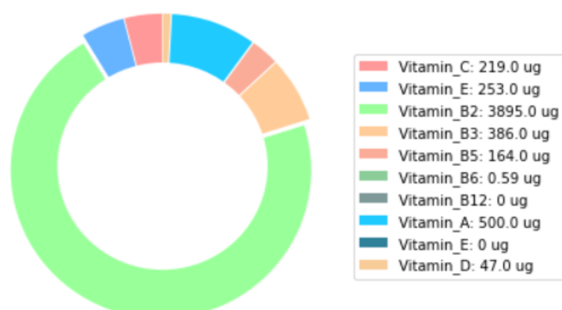
After training both the Inception V3 model and the ResNet50 model, we compared their accuracy to decide which one to choose. The following picture shows the accuracy vs epoch for Inception V3 model and ResNet50 model among 32 epoches training. Inception V3 (acc 0.8701) shows a better accuracy than ResNet50 (acc 0.8135) in our result. So we chose the model Inception V3 as our final model. (Detailed training process can be checked in Resnet50.ipynb and 3303FinalProjectTrainIncV3.ipynb)

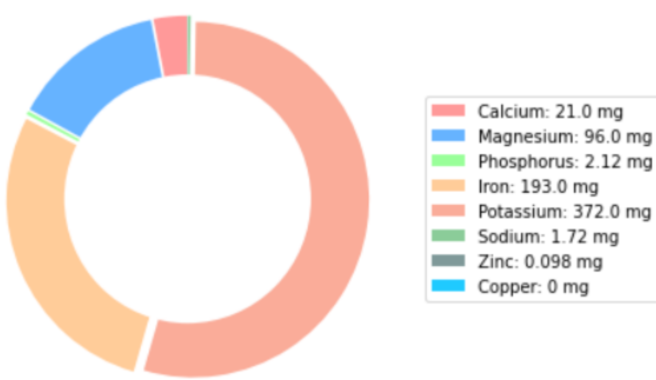



For a user, he can test our model in Demo.ipynb. After importing the Excel file we produced before and the final model we trained (both can be found on github), he can upload the food image to get the name of food and its nutrition details. Below are the figures that show the outputs of our project. The left side figure shows the name of food our model predicted and also shows the food image the user upload. The figure located at the right side and the figure below it show the nutrition details of the input food, such as its energy, its nutrition score, and show the vitamin percentage and the mineral percentage in form of pie charts.

```
Energy:  226.0 kcal/100g
Score:  C
```

```
hamburger
```



## Observations

Choosing the right model is important. Using the same input dataset, different models had different accuracy. We trained ResNet50 at first, and we got our "final model" with accuracy of 0.8135. Then we trained the Inception V3 model and we got our new final model with accuracy of 0.8701. Therefore, by changing the model, we improved the accuracy of our final model from 0.8135 to 0.8701. And the lower the learning rate, the higher the accuracy. We trained the Inception V3 model two times with different learning rates and we found the one with lower learning rate had higher accuracy. In addition, for the same model, different food types had different accuracy. When only testing the model with the images of popular food, it had relatively higher accuracy than only testing the images of unpopular food. For example, When only testing the images of donuts, our final model had accuracy around 0.6. When only testing the images of waffles, our final model had accuracy around 0.3, which was lower than the donuts case. (Detailed can be checked from PredictOneKind.ipnyb)

## Conclusion

In our project, the Inception V3 shows a better accuracy than ResNet50. Adjusting hyperparameters and choosing models are both important.

Also, it's hard to define a kind of food. When we collected the data, we found that if you search for a kind of food, there will be hundreds of products shown as results. Take humburger as an example, it can be a beef burger, egg burger or even a veggie burger. It's hard to give a comprehensive definition whether the food is healthy or not. In our project, we just choose the general case. In the future work, we need more specific classification which means larger quantities of labeling. In addition, our project can only output the nutrition details for a specific type of food. If there is an image with many kinds of food, for example, a hamburger and a french fries on one image, we can't output the nutrition details for both foods at one time. Therefore, the way we can improve is to consider incorporating food detection into our project.

## Reference


1.Happyforks.com

https://happyforks.com/

2.Open food facts dataset

https://world.openfoodfacts.org/

3.Snap-n-eat github repo

https://github.com/gabrielilharco/snap-n-eat

4. Food Image Classification with Convolutional Neural Networks,  Malina Jiang
http://cs230.stanford.edu/projects_fall_2019/reports/26233496.pdf

5.Convolutional neural networks are fantastic for visual recognition tasks
https://cv-tricks.com/cnn/understand-resnet-alexnet-vgg-inception/