Design and Development of Mobile App for Food Recognition

Chandrasekaran.G, Murugachandravel.J, Neethidevan.V, Karthikeyan.A

Abstract— Obesity is one of the major problems many of the people are facing today. Obesity occurs because of excessive intake of food and lesser physical exercise or no physical exercise. A limited number of cases are due to genetics, medical reasons, or psychiatric illness. In order to have a balanced diet, we need to track the food calories, proteins, and minerals. It is hard to search each and every time about nutritional facts and also it is time consuming. In "Food Lens app", simple snap of food photo will pop up all nutritional facts about the food and will be displayed in a user-friendly manner. Food Lens app uses modern Deep Learning methods to predict the food and related nutritional facts mapped in a faster manner. Now this app is able to predict ten different continental foods which have excess calories. So, this app prevents the users from taking excess calorie food or would assist them to take proper food periodically.

Keywords: Obesity, Food lens, Convolutional Neural Network (CNN)

1. INTRODUCTION

Everyone would like to have a healthy and tasty food. A hygienic and healthy food provides us with vital nutrients, like, proteins, vitamins, carbohydrates, fluids and fibres. Nevertheless, today, most of the young people are following diet pattern followed in Western countries. These patterns include meat, milk products, and artificially processed foods, lesser intake of fruits and vegetables, fish, legumes and grains. Keeping these problems in mind, Food Lens app helps the user to know about exact nutritional fact of the food before eating it, thereby preventing the people from taking food with excess calorie. The user has to simply snap the food and he/she can get the nutritional information about the food. Food Lens App is implemented by the Food Artificial Intelligence (AI) model. Food AI Model is based on the deep learning and image classification technology to identify food with much ease.

2. LITERATURE SURVEY

As human beings, we choose our food based on our own knowledge and experience. Machines use databases to store details about food that can be queried at any time for information. Man seeks the advices of machines in choosing his food and be under control. Neil Grimmer, Habit CEO, explains that, as human beings, we all know which food is good and healthy for us. He also points to a world where some people follow the food pyramids and generic advice. Raza Yunus et al., [1] proposed a method using deep learning to estimate food attributes including its ingredients and nutritional value by classifying the input food image. Parisa Pouladzadeh et al., [2] proposed a method for determining food calorie and nutrition that can help patients and dietitians to know about the amount of food they can consume daily. Kiran Ambhore et al., [3] proposed a method for classifying the food based on colour, shape, size and texture using K – nearest neighbor (KNN) classifier.

3. IMPLEMENTATION AND RESULT

System Implementation

The system is implemented via the following steps:

- Dataset Collection
- Dataset Pre-processing
- Model Construction
- Training
- Testing
- Deployment of the Model

3.1 Dataset Collection

For the food lens app, food images were collected. The dataset contains different types of food. A simple training set for image analysis was made. The data has been downscaled, so that test on the images can be made simpler and quicker. The data has been reformatted as **Hierarchical Data Format (HDF5)** so that the images can be easily read in. The file name indicates the information about the file. For example. food_c10_n1000_r384x384x3.h5 means that there are 10 categories represented, with n=1000 images, that have a resolution of 384x384x3 (RGB, uint8). Food_test_c10_n1000_r32x32x1.h5 means the data is part of the validation set, has 10 categories represented, with n=1000 images, each having a resolution of 32x32x1 (float32 from -1 to 1)

Categories of foods include 1.Chocolate cake, 2.Donuts, 3.French fries, 4.Fried rice, 5.Hamburger, 6.Ice cream, 7.Omelette, 8.Pizza, 9.Samosa, 10.Spring rolls

Overview of Dataset

Source	No. Of Training Image per Class	No. Of Testing Image per Class
Kaggle	700	300
Google	750	100
Total	1450	400

Revised Version Manuscript Received on April 05, 2019.

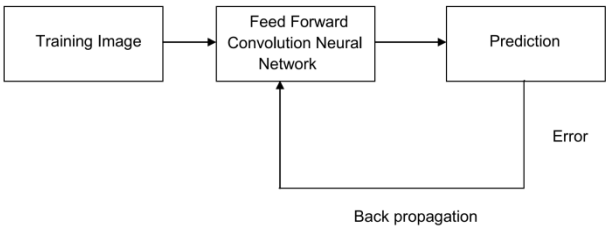
Chandrasekaran.G, MCA Department, Mepco Schlenk Engineering College, Sivakasi, Tamil Nadu, India
Murugachandravel.J, MCA Department, Mepco Schlenk Engineering College, Sivakasi, Tamil Nadu, India
Neethidevan.V, MCA Department, Mepco Schlenk Engineering College, Sivakasi, Tamil Nadu, India
Karthikeyan.A, MCA Department, Mepco Schlenk Engineering College, Sivakasi, Tamil Nadu, India

Number of categories of Food - 10
Total Training Image (1450 * 10) - 14500
Total Testing Image (400 * 10) - 4000

3.2 Dataset Preprocessing & Simulation Results

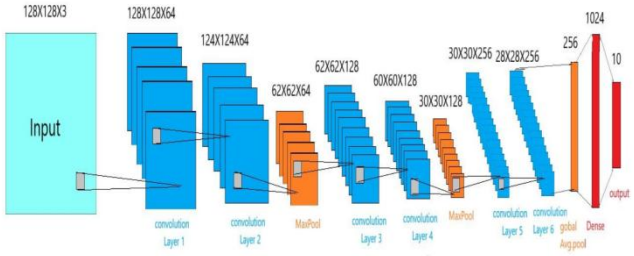
From the meta data available in "train.txt", the path of the file is read using pandas library. TRAIN_PATH = 'meta/' TRAIN_IMG_PATH = "images/" f = open(TRAIN_PATH+"train.txt","r") lists = f.readlines(). The image was read as coloured image and rearranged from BGR to RGB. Then the image was resized to 128x128 resolutions. If needed, the image can be resized using the OpenCV function cv2.resize(). The size of the image can also be specified manually, or by specifying the scaling factor. Different interpolation methods like cv2.INTER_AREA are used for shrinking and cv2.INTER_CUBIC (slow) & cv2.INTER_LINEAR are used for zooming. The default interpolation method used for resizing is cv2.INTER_LINEAR.

3.3 Model Construction



In this system, the feed forward CNN used on the training image. If the predicted output is not correct, back propagation is used.

Convolution Operation:



In the above diagram, width, height and depth of the given image are considered. The depth of the image represents the number of filters used. There are 8 layers in the CNN. Layers 1 to 6 act as a sparse network. A convolution refers to the procedure where two sources of information are intertwined. A kernel or a filter, is a 3 x 3 matrix, with a stride 1. In all maxpool layers, 2 x 2 matrix is applied with a stride 2. Layers 7 and 8 are fully connected and layer 8 acts as the output layer having 10 neurons, each showing the probability of the food. Kernels are convolved with the given image to obtain, an activation maps or feature maps. Activation maps represent the 'activated regions', ie., regions which contain the features specific to the kernel in the input image. The value of the kernel matrix changes with every learning iteration over the training set indicating that the network is learning to identify which regions are appropriate for extracting features from the data. Next, dot product between the kernel and the input matrix was computed. A single value in the activation matrix is the convolved value obtained by summing the terms obtained from the dot product of the activation matrix. The selection of patches moves towards right or down when the boundary of the matrix is reached by a stride value. This process is repeated until entire image is processed. The same process is repeated for all the three colors channels. Usually, every neuron will be connected to every pixel. Nevertheless, we describe 2D region, of size 5 x 5 units and then extend over for the three colors also. All the pixels are connected to the input layer of the neural network. The activation maps are produced by traversing the entire network layer cross section.

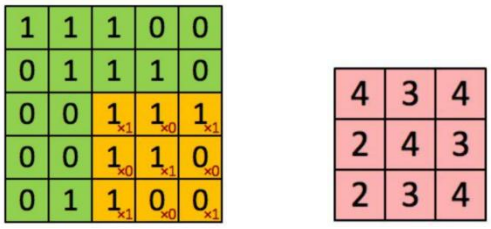
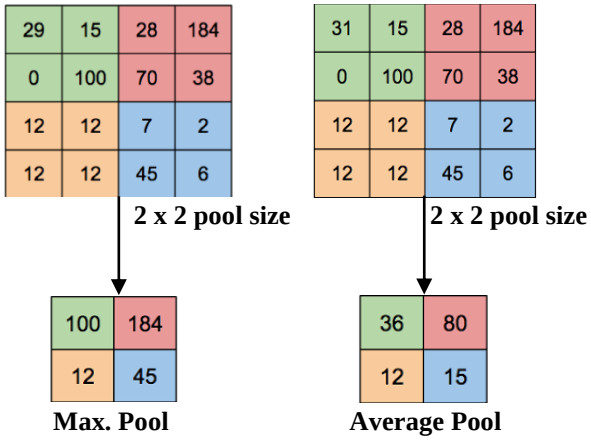


Image **Convolved feature**

To reduce the spatial dimensions of the input image, Pooling is done. The transformation of the image is performed by considering the maximum value or the average value of the window pixels. This process of transformation is referred to as down sampling. As Max pooling gives better performance, it is always given more preference.



Pooling is done to reduce the spatial dimensions of the input data to the next convolutional layer. The transformation is either performed by taking the maximum value or the average value from the window. It is also called down sampling. Max pooling is preferred due to its better performance. Max pooling is followed for first few layers and average pooling is followed for the last layer.

3.4 Training the Model:

While training a model, the model is given a known dataset. Called a training dataset, and an unknown dataset, called the validation dataset or testing set. The cross-validation is done to test the model's ability. Under each round of cross-validation, the sample data is partitioned into subsets, and each subset is analysed, called the training set.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

Validation of the analysis of the other subset is called validating set.

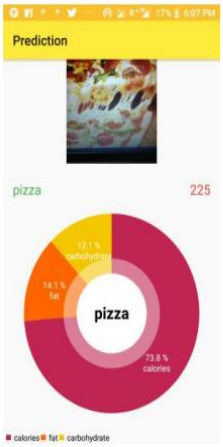
3.5 Deployment of Model:

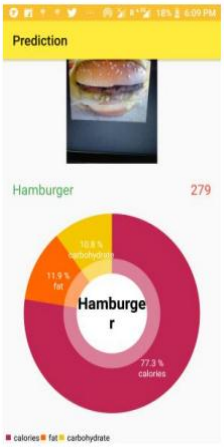
Camera Module:

This module is used to scan the food. It has click button for capturing the image and predicting the image. It also has the feature of front & rear camera and flashlight.

Output Module:

This module displays name and the nutritional facts about the food. It displays in the pie chart format. Response will be in the form of Json. By using Gson library, the json data is converted into Required Java object format.





4. CONCLUSION

By the Food Lens app, people can get quick attention about what they are eating and helping them to prevent unnecessary calories. By the time, Food lens app correctly predicts 10 different continental foods. In future, this app will train on more number of foods from all over the world and is the best food identification system on the market. Food Model accuracy will improves as new food images are added to the server.

REFERENCES

1. Raza Yunus, Omar Arif, Hammad Afzal, Muhammad Faisal Amjad, "A Framework to Estimate the Nutritional Value of Food in Real Time Using Deep Learning Techniques", Article in IEEE Access, December, 2018
2. Parisa Pouladzadeh, Shervin Shirmohammadi, Rana Al-Maghrabi, "Measuring Calorie and Nutrition from Food Image, "IEEE Transactions on Instrumentation and Measurement, Vol.63, Issue: 8, August 2014.
3. Kiran Ambhore, N.A. Dawande, "Measuring Calories and Nutrition from Food Image", International Journal of Advanced Research in Computer and Communication Engineering, Vol.5, Issue 6, June 2016
4. G. A. Bray and C. Bouchard, Handbook of Obesity, 2nd ed.BatonRouge, LA, USA: Pennington Biomedical Research Center, 2004.
5. Robert M. Haralick, K. Shanmugam, I. Dinstein, "Textural Features of Image Classification", IEEE Transactions on System, Man and Cybernetics, vol. 3, no.6, 1973
6. A. Jain and G. Healey, "A Multiscale Representation Including Opponent Color Features for Texture Recognition", IEEE Transactions on Image Processing, Vol. 7, No.1, January 1998.