International Journal of Science and Research (IJSR)
ISSN: 2319-7064
SJIF (2019): 7.583

Detection and Classification of Food Consumption Using Convolutional Neural Networks

Diksha Solanki

Department of Information Technology, Maharaja Agrasen Institute of Technology, Delhi, India

Abstract: ***Food monitoring and nutritional analysis assumes a main part in health related issues, it is getting more essential in our everyday lives. In this paper, we apply a convolutional neural network (CNN) to the task of detecting and recognizing food pictures. In light of the wide variety of kinds of food, image recognition of food items is commonly extremely troublesome. In any case, deep learning has been demonstrated as of late to be an extremely ground-breaking image recognition technique, and CNN is a best in class way to deal with deep learning. We applied CNN to the undertakings of food detection and recognition through boundary enhancement. Highlights learned by Convolutional Neural Networks (CNNs) have been perceived to be more robust and expressive than hand-created ones. They have been effectively utilized in various PC vision tasks, for example, object discovery, pattern recognition and picture understanding.***

Keywords: Food detection, Convolutional Neural Networks, Food recognition, Food Classification, Deep Learning

1. Introduction

Diet is significant in human life. Getting satisfactory nourishment from daily meals is basic for our wellbeing.

In the flow age, individuals are more cognizant about their food and diet to keep away from either forthcoming or existing infections. Since individuals are reliant on shrewd advances, arrangement of an application to naturally screen the people diet, helps in numerous perspectives. It expands the attention to individuals in their food propensities and diet. In the course of the most recent twenty years, research has been centered around consequently perceiving the food and their wholesome data from pictures caught utilizing PC vision and AI procedures. To appropriately survey dietary admission, precise assessment of calorie estimation of food is of principal significance. A greater part of individuals are indulging and not being sufficiently dynamic. Given how occupied and focused on individuals are today, it's easy to neglect to monitor the food that they eat. Likewise, considering the current situation of lockdown and isolation during the far and wide COVID-19 pandemic, numerous individuals are inclined to gorging and not dealing with their eating regimen. This lone builds the significance of proper classification of food.

A portion of the techniques presently being used for dietary appraisal include self-reporting and physically recorded instruments. The issue with such strategies for appraisal is that the assessment of calorie utilization by a member is inclined to bias, for example thinking little of and under detailing of food admission. To build the exactness and decrease the bias, improvements to the current strategies are required. One such potential arrangement is a portable distributed computing framework, which utilizes gadgets, for example, cell phones to catch dietary and calorie data. The subsequent stage is to naturally break down the dietary and calorie data utilizing the figuring limit of the cloud for a goal appraisal. In any case, clients actually need to enter the data physically. In the course of the most recent couple of years, a lot of innovative work endeavors have been made in the field of visual-based dietary and calorie data investigation. In any case, the effective extraction of data from food pictures stays a difficult issue.

Our problem consists of classifying images of food to three different classes. There are two main challenges in this task. Primary problem is that the same food is made differently depending on the location, the available ingredients and the personal taste of the cook.

Another one could be the angle from which the image was taken. These challenges cause significant variations within images from the same class and makes classification problem more difficult.

We chose to classify images of food since our intent was to make a program that could be used in dietary assessment applications. In this paper, we applied convolutional neural Networks (CNN) to tackle this issue. We will introduce the usage of our model just as the preparation cycle and results. In this paper, an exertion has been made to characterize the pictures of food items for additional eating routine observing applications utilizing convolutional neural networks (CNNs). Since the CNNs are equipped for dealing with a lot of information and can assess the features naturally, they have been used for the undertaking of food classification. The standard Food-101 dataset has been chosen as the working information base for this project.

2. Literature Review

Obesity is yielding a concerning issue in the present life. The transcendent explanation of weight is devouring a greater number of calories than we consume which can truly subvert the personal satisfaction. Researchers says, precisely evaluating dietary admission is a significant factor to decrease this danger. To meet this exigency, researches have adopted a few strategies to gauge the calorie of a food. In 2009, an extensive food image and video dataset was built named the Pittsburgh Fast-food Image Dataset (PFID), containing 4545 still images of 101 different food items, such as chicken nuggets and cheese pizza etc. The researcher had applied Support Vector Machine (SVM) classifier on this

International Journal of Science and Research (IJSR)
ISSN: 2319-7064
SJIF (2019): 7.583

dataset and had achieved a classification accuracy of 11% with the color histogram method and 24% with the bag-of-SIFT (Scale-Invariant Feature Transform)-features method. Chen et al. (2012) focused on this major issue and proposed a method of food detection and quantity estimation for the purpose of dietary analysis. They made use of Gabor and color features to represent food items. A multi-label SVM classifier combined with a multi-class Ad boost algorithm was used to represent that the new technique can improve the performance of original SIFT and LBP feature descriptors. Around 50 categories of food like soup, dumplings etc. were used and achieved 68.3% accuracy. Probst et al. (2015) is motivated to introduce another prototype for dietary analysis with the aid of smart phone as well as the features of image processing and pattern recognition. Scale invariant feature transformation (SIFT), local binary patterns (LBP), color etc. are common visual features that are used for food images. The bag-of-words (BoW) model was used to perceive the photos taken by the phone.

Deep learning gradually becomes a very important image recognition technique, and CNN is the most popular deep learning architecture. In 2015, Yanai et al. applied deep convolutional neural network (DCNN) technique on Image Net dataset and achieved an accuracy of 78.77% for UEC-FOOD100 and an accuracy of 67.57% for UECFOOD256 dataset. Kagaya et al., also applied CNN on their own dataset for the detection and recognition of the food. CNN provides higher accuracy than usual support-vector-machine-based methods where the accuracy rate for recognition was 73.70% and for detection was 93.80% . In 2016, Hassannejad et al. proposed a deep convolutional neural network (DCNN) technique which had a depth of 54 layers on UEC FOOD 100, ETH Food-101 and UEC FOOD 256 dataset and he achieved 88.28%, 76.17% and 81.45% as top-1 accuracy and 97.27%, 96.88% and 92.58% as top-5 accuracy for dietary analysis. Christodoulidis et al. applied a 6-layer deep convolutional neural network (CNN) on their own dataset containing 573 food items to classify food and the accuracy rate was 84.9%. In 2016, Singla et al., proposed a new method of identifying food/non-food items and recognizing food successfully using a GoogLeNet model based on deep convolutional neural network. According to their experimental results, a high accuracy rate of 99.2% was achieved in food/non-food item classification and 83.6% respectively in food item recognition. Liu et al., also proposed a new CNN based food recognition algorithm and applied it on UEC-256 and Food-101 data sets and achieved 87.2% and 94.8% accuracy respectively. In a five layer CNN with bag-of-features (BoF) and support vector machine was applied on a dataset containing 5822 images of ten categories with the overall accuracy of 56%. After that data expansion techniques was applied to increase the size of training images for which the accuracy increased by 90%.

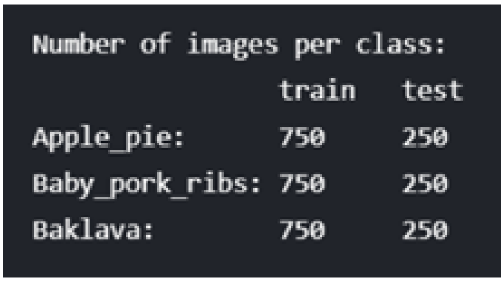
Due to the complexity of food images, many of the earlier proposed methods for food recognition achieved low classification accuracy. In our proposed system we used the FOOD 101 training data set that is publicly available data set. We use trained CNN to extract and to classify food images of ten different classes and achieved accuracy 78.9%.

2.1 Dataset

For our purposes, we selected 3 classes from challenging Food-101 dataset, because we considered them to have the most distinct and representative colours and patterns.

This dataset consists of 101 categories and each category has 1000 images, in total 101,000 images. Each one only contains the label information indicating the food type. Most of the images are popular western food images. They haven't been taken in laboratory conditions, but are randomly taken pictures by people around the globe. Each food class contains 1000 images. Using the data provided, a deep learning model built on Keras/Tensor Flow is trained to classify 3 classes in Food 101 dataset.

The 3 different classes of food chosen are:
1) Apple Pie
2) Baby Pork
3) Baklava


Number of images per class:

	train	test
Apple_pie:	750	250
Baby_pork_ribs:	750	250
Baklava:	750	250

Figure 1: Dataset

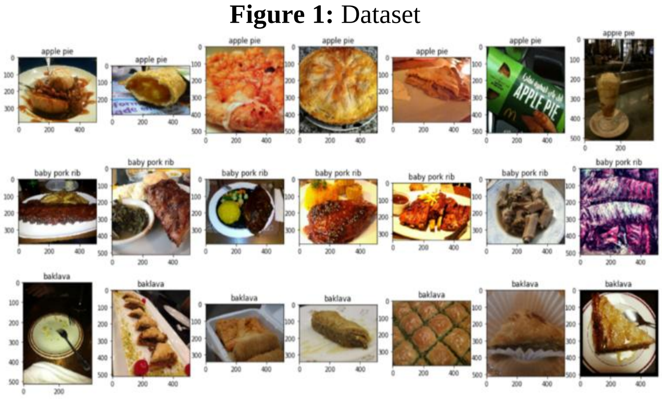


Figure 2: Dataset Images

3. Methods

Before describing the architecture and the different components of the proposed system, we provide a brief introduction to the deep CNNs.

3.1 Convolutional Neural Networks (CNN)

The CNN offers a cutting edge strategy for image recognition. It is a multilayer neural organization, whose neurons take little fixes of the past layer as information. It is hearty against little moves and revolutions. A CNN framework includes a convolution layer and a pooling (or sub-sampling) layer. In the convolution layer, not at all like for general completely associated neural organizations, weights can be considered as n × (n < input size) channels. Each information convolves these channels. Each layer has numerous channels that create various yields. For the image recognition task, the various highlights are separated by these channels. The channels are regularly called (convolution) parts. The pooling layer delivers the yields by actuation over rectangular areas. There are a few actuation strategies, for example, maximum and average activation. This makes the CNN's yields more invariant with regard to position. A commonplace CNN contains various convolution and pooling layers, with a completely associated layer to create the eventual outcome of the undertaking. In image recognition, every unit of the last layer shows the class likelihood. A CNN has hyper boundaries that incorporate the quantity of center layers, the size of the convolution parts, and the dynamic capacities.

4. Methodology

At the earliest reference point of our test technique, it is important to do a few preprocessing to prepare the pictures for work appropriately. Fig 3 shows the total technique of our proposed framework.

Our model is actualized as a convolutional neural organization that comprises of three convolutional layers, continued by ReLU actuation and max pooling and, a solitary completely associated layer and a softmax layer, as seen on Figure 3. We utilized group standardization after every activation function, as it improved the accuracy of our model by around 2%. As an optimizer, we utilized Adam with default learning pace of 0.001, as it beat other improvement calculations (Adagrad, SGD). Cross entropy was utilized as a loss function.

4.1 Preprocessing

A raw picture contains of specific factors, for example, commotion, climatic conditions, bad resolution and undesirable background for which it isn't reasonable enough to characterization and recognizable proof. So it is essential to improve picture quality and set up the picture for additional preparing to recognize the item as precisely as could be expected under the circumstances. In this paper the pre-preparing measure comprises of noise decrement and contrast enhancement.

4.2 Preparing Training and Test image sets

We have separated the whole dataset into two subsets specifically the training set and validation or testing dataset. 30% pictures were arbitrarily chosen for training dataset and the rest of pictures for test datasets. Our dataset is made up by three unique sorts of food namely apple pie, baby pork ribs and baklava.

4.3 Train the CNN Classifier

Convolution neural network, a generally utilized deep learning tool are enlivened from the natural structure of a visual cortex. Alongside information and yield layer CNN comprises of various concealed layers, for example, convolutional layers followed by max-pooling layers, and fully connected layers. The design of a convolution neural organizations can fluctuate contingent upon the sorts and numbers of layers included.

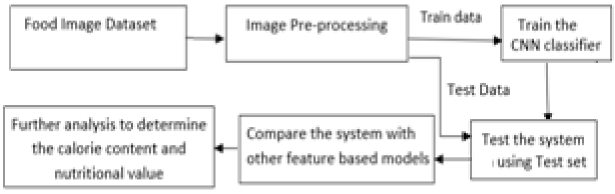


Figure 3: Block diagram of the proposed methodology

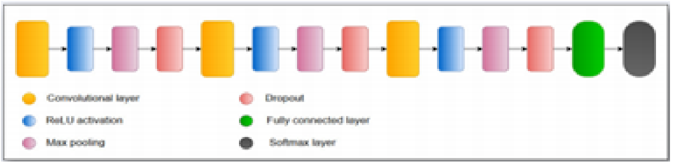
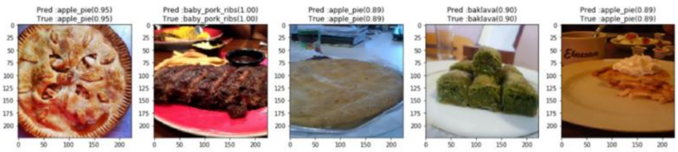


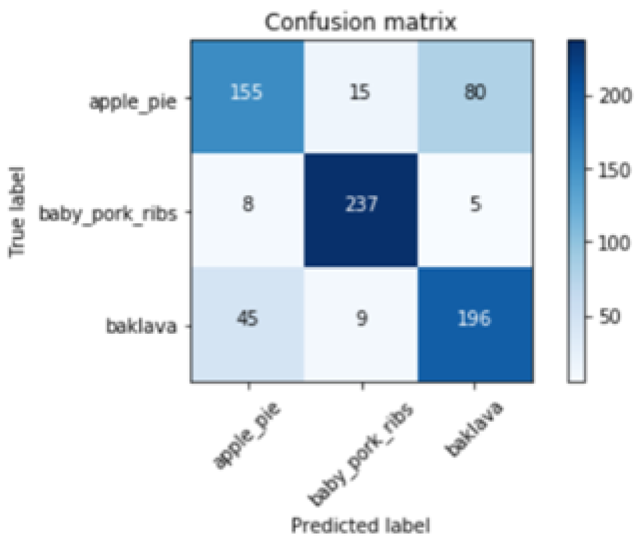
Figure 4: Model Architecture

5. Experimental Results

We divided our dataset into three parts: training (70%), validation (10%) and testing (20%). Model was trained for 100 epochs with batch size of 64 (each epoch having 20 steps). This model has the peak validation accuracy of 78.9% on classifying images to three classes (apple pie, baby pork, baklava). Below are predictions from the model:



The confusion matrix of 750 test images:



6. Conclusions

This project shows that convolutional neural network can successfully solve food image classification problems to relatively small number of classes. Classification to more classes requires more complex architectures. Besides the complexity of a model, it is also important to choose adequate hyper-parameters and optimization functions. As shown in the confusion matrix, most of the wrong prediction are between apple pie and baklava. The performance of the

Paper ID: SR201130135219 DOI: 10.21275/SR201130135219

apple pie and baklava are not as good, this might be explained by that fact that both of these food types have similar texture and color, as both are made from pastry and the model finds it harder to classify between them.

With the given data sets for 3 classes of food: apple pie, baby pork ribs and baklavas, the model final accuracy reached 78.9%. The main cause of error is due to the similarity between baklavas and apple pie as they both exhibit alike texture and colors.

While solving this problem, we came to a conclusion that image augmentation is of great importance when using complex datasets. Image augmentation also helped us deal with over fitting. It decreased fluctuations in validation accuracy/loss and allowed them to increase/decrease more steadily here.

7. Acknowledgement

I would like to express my gratitude towards my mentors Dr. Amita Goel, Ms. Vasudha Bahl and Ms. Nidhi Sengar for their guidance and constant supervision as well as for providing necessary information regarding the project and also for their support in completing the project.

References

[1] Dixon, J. B. (2010). The effect of obesity on health outcomes. Molecular and cellular endocrinology.
[2] Chen, M.; Dhingra, K.; Wu, W.; Yang, L.; Sukthankar, R.; Yang, J. PFID: Pittsburgh Fast Food Image Dataset. In Proceedings of the ICIP 2009, Cairo, Egypt, 7–10 November 2009;
[3] Chen, M. Y., Yang, Y. H., Ho, C. J., Wang, S. H., Liu, S. M., Chang, E., & Ouhyoung, M. (2012, November). Automatic chinese food identification and quantity estimation. In SIGGRAPH Asia 2012 Technical Briefs (p. 29). ACM.
[4] Lowe, D.G. Object Recognition from Local ScaleInvariant Features. In Proceedings of the ICCV'99,Corfu, Greece, 20–21 September 1999;
[5] Yanai, K., & Kawano, Y. (2015, June). Food image recognition using deep convolutional network with pre-training and fine-tuning. In Multimedia & Expo Workshops (ICMEW), 2015 IEEE International Conference on (pp. 1-6). IEEE.
[6] Kagaya, H.; Aizawa, K.; Ogawa, M. Food Detection and Recognition using Convolutional Neural Network. In Proceedings of the MM'14, Orlando, FL, USA, 3–7 November 2014; pp. 1055– 1088.
[7] Probst, Y., Nguyen, D. T., Tran, M. K., & Li, W. (2015). Dietary assessment on a mobile phone using image processing and pattern recognition techniques: Algorithm design and system prototyping. Nutrients, 7(8), 6128-6138.