# DS 863 : MACHINE PERCEPTION

## PROJECT REPORT

**Submitted to:**

Prof. Dinesh Babu J.

**Submitted by:**

Arhima Singh (MT2016022)

Dheeraj Doodhya (MT2016052)

Pritesh Mandwariya (MT2016110)

Rajat Batra (MT2016114)

Group #G19

**Date of Submission:**

16 May 2017

# INDEX

# 1. TITLE

Restaurant Famous For (RFF)

# 2. ABSTRACT

Food nowadays is considered to be one of the most photographed objects. This is evident from the abundance of food photography in social networks, photo sharing websites and mobile applications. In the age of selfies and photo-centric social storytelling, web users upload an enormous amount of photos of different food items and different food restaurants everyday alongside their written reviews.

In this project, we are trying to analyze this data set of food images uploaded by the user on social networking websites and trying to figure out, what food item a particular restaurant is famous for. We experimented with different recognition methods and assessed their effectiveness in food classification.

# 3. DESCRIPTION OF PROBLEM AND PROJECT OBJECTIVE

Today we have social networking websites (like zomato, instagram etc.) where users upload the images of food items whenever they visit some restaurant, cafe, bar, hotel etc. Most of these websites provide rating of these places on some scale and what cuisine is served over there. The rating is based on the users experience on visiting the place. However, no social networking website, based on users experience, tells what that restaurant/cafe/hotel is famous for, what food item most people order on visiting that place. Though the websites (like zomato.com) specifies the cuisines they serve but this is based solely on what restaurant owners mention on registering itself to the website and does not involve any user review.

With this project, we tried to achieve following objectives:

1. Construct models for doing classification on available Food-101 dataset of food images. Obtain the most effective classification model which yields the highest performance on food images.

2. Based on the dataset of food items (Food-101 dataset) from the social networking websites, we would tell what a particular restaurant / cafe / hotel is famous for.

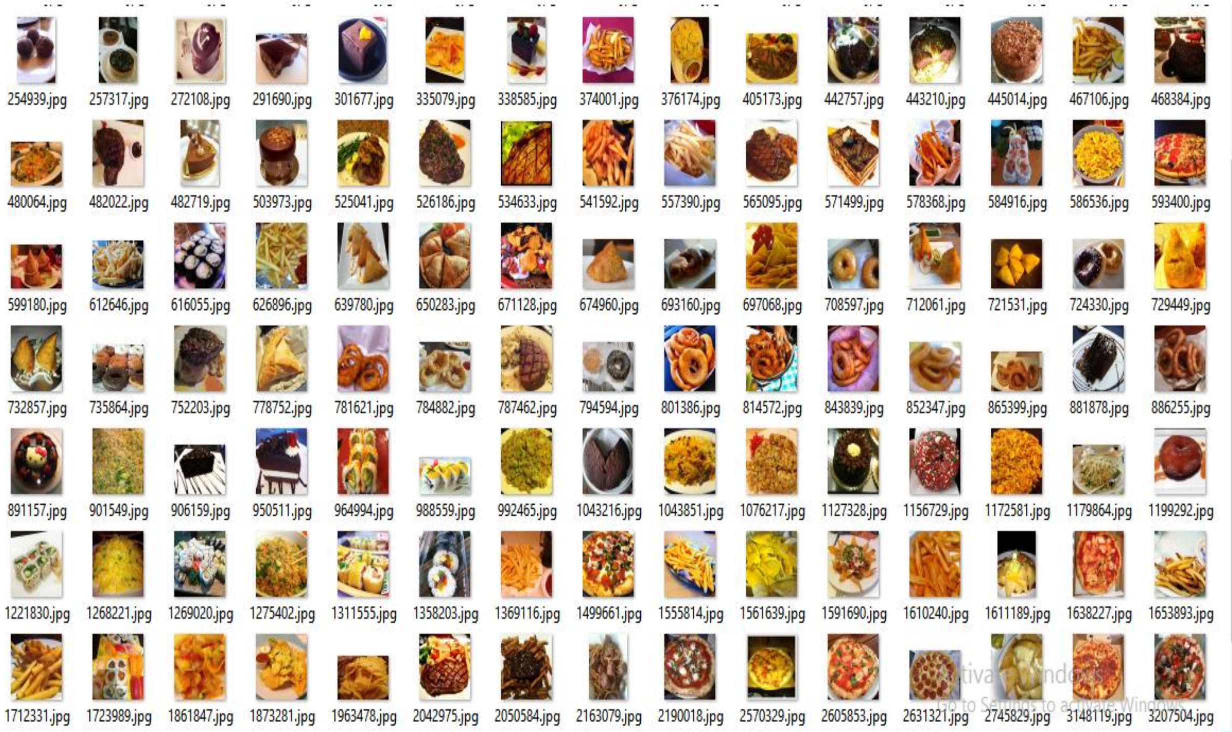
# 4. DATA SET USED

We are using a data set which contains the images of food items called 'Food-101'. Food-101 data set was created by Computer Vision Laboratory, Zurich, Switzerland. This dataset is made up of real world food images downloaded from foodspotting.com , a site which allows users to take pictures of what they are eating, annotate the image with label or caption the image and upload the image online. It consists of 101 different food categories and there are 1000 images present for each food item. So, instead of working on standardized food images, we'll be working on real time food images which would be having irregularities and variations and so our project would be based on real time images and not only on standard images of food items.

We'll be using a subset of this dataset and would take into consideration only a few food items. The food categories we considered are:

| | | | | | |
|---|---|---|---|---|---|
| Chocolate cake | Sandwich | Donuts | French fries | Nachos | Onion rings |
| Pizza | Samosa | Spaghetti | Steak | Sushi | Fried Rice |

Sample food images from selected 12 categories are shown below:



# 5. APPROACHES USED

To identify what food the restaurant/cafe/hotel is famous for, first we need to perform food item identification for the food images uploaded by the user. To perform classification of food images under different sections, we tried different approaches. First we tried using feature extraction methods, SIFT and SURF, and got reasonable accuracy using these techniques. Then, in order to increase the accuracy of classification we researched about other ways of classification and tried a few more approaches. We tried multiclass classification using SVM, with linear Kernel and RBF-Kernel. Of these two approaches, as expected, SVM with RBF kernel gave better accuracy. Results achieved with SVM using RBF-Kernel were better as compared to SIFT and SURF. Highest accuracy was achieved when we used LP with AdaBoost. As a result, the final implementation of our project is based on classification with LBP using AdaBoost.

## Approach 1: Using LBP with AdaBoost

**Local Binary Patterns:**

LBP is a powerful operator of texture description. It labels the pixel of an image region by thresholding the 3 X 3 neighbourhood of each pixel with the centre value and by arranging the result to binary codes.

ex. Let a pixel and its neighbouring pixel values are as shown:

| | | |
|---|---|---|
| 80 | 76 | 65 |
| 79 | 75 | 70 |
| 69 | 74 | 80 |

| | | |
|---|---|---|
| 1 | 1 | 0 |
| 1 | | 0 |
| 0 | 0 | 1 |

LBP pattern corresponding to given pixel and its neighbouring values is (10010011). We compute a 256 bin histogram of the labels which can be used as a texture descriptor. This histogram statistically describes the image characteristics. Each bin of histogram can be regarded as a micro-texton. Local primitive patterns which are codified by these bins include different types of curved edges, spots, flat areas etc. The image is equally divided into smaller regions to extract LBP Histogram. The LBP features extracted from each sub-region are concatenated into a single spatially enhanced histogram. Each histogram would contain 59 bins, corresponding to 59 uniform patterns. Bin to bin distance is used to compare histograms.

**LBP with AdaBoost:**

AdaBoost learns a small number of weak classifiers and boosts them iteratively into strong classifier of higher accuracy. At each iteration, a weak classifier which minimizes the error rate is selected and the distribution is updated to increase the weights of misclassified samples and reduce the importance of others. AdaBoost is actually used to find the sub region that contains more discriminative information.

Also, AdaBoost works for 2 classes, however, in our project, the food items can create K different classes. So, multiclass classification problem is accomplished using the one-against-rest technique, which trains AdaBoost between one expression with all others.

**Implementation steps** of this approach are described in following steps :

**1. Loading training data set :** The training data set consists of different packages with names of food items labelled on them. Each of these packages contains 1000 images of that food item. A random 70% of this dataset would be loaded as training images.

**2. Loading test data set :** The remaining 30% of the images in the dataset would be loaded as test images. These images correspond to the images of food items from different restaurants / cafes / hotels.

**3. Creating a Histogram of training dataset using LBP:** We iterate over all the images in the training dataset and a histogram is created for images in the training dataset. All images are resized to a fixed size of 500 X 500 and are converted to greyscale. While iterating over images, each image is divided into 20 sub regions, each of size 125 X 100.

Local binary patterns are obtained for each subregion using the function *feature.local_binary_pattern()*. This function takes number of circularly symmetric neighbour set points considered for each image and the radius of the circle from the pixel. We also pass a parameter 'nri_uniform', which corresponds to non rotation-invariant uniform patterns variant. This makes the LBP pattern invariant to rotation of object. The information in the bins are converted to a 1-D array using the function *numpy.arange()* and we create a histogram out of this array. For each sub-region, the LBP histograms in a given class are avargaed to generate a template for that class of food item.

**4. Fitting a model :** The model takes a form of histogram representation of the images, based on the collection of its local features. We then apply AdaBoost on the histogram and model is trained with the training data. This is done using function

*model = AdaBoostClassifier().*

**5. Creating Histogram for test dataset :** A histogram is created for the images in the test dataset in a way similar to step 3. This histogram is compared with the histogram of training images and difference between the two histograms is calculated. Labels are predicted for the images in the test dataset.

*predictedLabels=model.predict(test_data)*

These predicted labels are compared to the actual values of the labels of the test images to calculate the accuracy of the predictions.

**6. Finding the popular food :** Once the labels are predicted, we find out the labels which appears maximum number of times. This label is returned as food the restaurant is famous for. Let the label with max frequency constitutes to x % of all images. We set a threshold of 10% such that all food with frequency greater than or equal to (x-10)% would be returned as the famous foods as well.

Apart from performing the classification of images with LBP and AdaBoost, we tried several other approaches which however, gave lower accuracy as compared to LBP with AdaBoost. Other approaches tried by us are discussed below:

## Approach 2: Using SIFT and SURF

With SIFT and SURF, for any food item given in the food image, we detect the interest points which provides the "feature description" of the food item. This description extracted from the training images can then be used to identify the object when attempting to locate the food item in test images. Once we got all the features, K-Means algorithm is applied over the extracted features of the images in the training dataset to cluster the features. The value of 'K' which gives maximum accuracy is selected.

A histogram of the features generated from the images in the training dataset is created. Images are then compared and classified based on this histogram representation. Using logistic regression, we try to fit the model according to the histogram of training images. A histogram of test dataset is also created and using this histogram we predict the labels for the food images in the test dataset using logistic regression.

With above approach, the accuracy of prediction was not good enough, which led us to trying other approaches.

## Approach 3: Using Support Vector Machines

Support Vector Machines are primarily designed for 2-class classification. However, K-SVMs can be used to solve K-classification problem. SVM can be used with one-versus-rest technique to solve K-classification problem. K-classification SVM consists of a learning module and a classification module. The classification module can be used to apply the learned module to new examples. If we have K different classes, then (K*(K-1))/2 classifiers are constructed and each one trains data from two different classes.

In our project, we tried SVM with:

(i) Linear Kernel
(ii) RBF-Kernel

Applying SVM with RBF kernel gave more accuracy in classification of food images. The results of applying SVM on our dataset are discussed in 'Results' section.

## Approach 4: Applying SVM with AdaBoost

In this approach SVM is used as a weak-learner for AdaBoost. This algorithm adaptively adjusts the Kernel parameters in SVM instead of using a fixed one. However, when we tried using this approach, the accuracy of the classification came out to be lesser as compared to using AdaBoost with LBP.

Of all the approaches discussed above, the approach which gave best results is using LBP with AdaBoost. The accuracy results of using different approaches are discussed and compared in Results section. Hence, the final implementation of our project is done using LBP with AdaBoost.

# 6. CHALLENGES

Classification of food images under different categories is a challenging task due to high variability of food images and extracting all the features and matching them is not done very accurately by the commonly used algorithms (like SIRF, SURF. LBP etc.). Same food items at different places can vary in appearance, shape, color etc. Also, the images taken by the user can be noisy with varying illumination and contrast due to the surrounding light. The images taken by the user can be from varying camera angles and can have overlapping food items.

Example, in the figure below, all images would come under the category 'Pizza'. All these images have varying orientation, illumination, shape etc but all of them full under the same class.


All such factors, makes the classification of food component difficult. The classification system needs to use the most discriminative features from the available data. Also, the training dataset of the food items plays a crucial role in classifying the images under different sections and so should be large enough and should have variety of images.

# 7. RESULTS

We considered a dataset of 12000 images, 1000 images of 12 different food items. We considered 12 different food items : chocolate cake, sandwich, donuts, fried rice, french fries, nachos, onion rings, pizza, samosa, spaghetti, steak, sushi.

Of the given dataset, a random 70% of the images constitute the training data set and remaining 30% constitutes the test images.

Following table shows the accuracy of classification achieved with different approaches:

| Approaches | Accuracy (Approximate) |
|---|---|
| **LBP with AdaBoost** | 81% |
| **SIFT and SURF** | 57%(SIFT), 43%(SURF) |
| **Support Vector Machines With RBF kernel** | 68% |
| **SVM with AdaBoost** | 54% |

The highest accuracy is achieved when classification is performed using LBP with AdaBoost, which is 81.138888%.

Following is the output achieved on execution of the program:

```
In [2]: runfile('D:/Python/Project/projectCode.py', wdir='D:/Python/Project')


Total Training images :                 8400
Total Testing images :                  3600
Number of Images Predicted Correctly :  2921
Number of Images predicted Incorrectly :        679

 Accuracy of Prediction : 81.13888888888889 %


The categories of food items we consider in our project are :
1       chocolate_cake
2       sandwich
3       donuts
4       french_fries
5       fried_rice
6       nachos
7       onion_rings
8       pizza
9       samosa
10      spaghetti
11      steak
12      sushi


The Restaurant is famous for following food items :
chocolate_cake
```

# 8. CONCLUSION

In this project we applied different image classification techniques for understanding of food images. We tried different approaches: using feature extraction methods, SIFT and SURF, multiclass classification using SVM, with linear Kernel and RBF-Kernel, LBP with AdaBoost, SVM with AdaBoost. Our experiment with different techniques shows that using LBP with AdaBoost, we can achieve better classification results. Highest overall accuracy rate achieved using LBP with AdaBoost is 81.138888%.