Pizza vs Not Pizza Classification

MOHAMED TAREK ELSEBAIE

Pizza vs Not Pizza Classification

MOHAMED TAREK ELSEBAIE

**Pizza and Not Pizza Classification**

**Introduction:**

To detect the images received by using image processing which classifies objects and facilities in high-resolution. So, we use image processing to make some enhancement on pictures to classify between the two classes. Our dataset consists of two classes in each class there is 983 images, one class for pizza and the other for not pizza.

The task of detecting whether an image contains pizza is a binary classification problem namely pizza/not-pizza classification. Given an image a pizza classifier identifies an image as pizza or not-pizza. This is like an image classification problem where a classifier is trained on image data using machine learning techniques. Classical approaches to image classification extract features such as interest point descriptors from scale-invariant feature transform, pool feature into a vector representation and then use cluster algorithm such as Support vector machine for classification.

Image processing is performing operations on an image in order to enhance the image or to extract some useful information from it. Since machines do not directly understand what is in the images like human brains, we use image processing to get different features for each image after enhancing it by removing noises and unwanted data.

**Related Work:**

**• Mariam, I. and Hassan, G. “Feature extraction using median–mean and feature line embedding” Faculty of Electrical and Computer Engineering, Tarbiat Modarres University, Tehran, Iran in 2015**

Link of the paper: https://arxiv.org/ftp/arxiv/papers/1606/1606.02210.pdf

• In this paper, the authors proposed a feature extraction method based on median–mean and feature line embedding.

**• Food-101 – Mining Discriminative Components with Random Forests:**

In this paper we address the problem of automatically recognizing pictured dishes. To this end, we introduce a novel method to mine discriminative parts using Random Forests (rf), which allows us to mine for parts simultaneously for all classes and to share knowledge among them. To improve efficiency of mining and classification, we only consider patches that are aligned with image super pixels, which we call components. To measure the performance of our rf component mining for food recognition, we introduce a novel and challenging dataset of 101 food categories, with 101’000 images. With an average accuracy of 50.76%, our model outperforms alternative classification methods except for CNN, including svm classification on Improved Fisher Vectors and existing discriminative part-mining algorithms by 11.88% and 8.13%, respectively. On the challenging mit-Indoor dataset, our method compares nicely to other s-o-a component-based classification methods.

Link of the paper: https://data.vision.ee.ethz.ch/cvl/datasets\_extra/food-101/static/bossard\_eccv14\_food-101.pdf

Keywords: Image classification, Discriminative part mining, Random Forest, Food recognition.

**• Wide-Slice Residual Networks for Food Recognition:**

Food diary applications represent a tantalizing market. Such applications, based on image food recognition, opened to new challenges for computer vision and pattern recognition algorithms. Recent works in the field are focusing either on hand-crafted representations or on learning these by exploiting deep neural networks. Despite the success of such a last family of works, these generally exploit off-the shelf deep architectures to classify food dishes. Thus, the architectures are not cast to the specific problem. We believe that better results can be obtained if the deep architecture is defined with respect to an analysis of the food composition. Following such an intuition, this work introduces a new deep scheme that is designed to handle the food structure. Specifically, inspired by the recent success of residual deep network, we exploit such a learning scheme and introduce a slice convolution block to capture the vertical food layers. Outputs of the deep residual blocks are combined with the sliced convolution to produce the classification score for specific food categories. To evaluate our proposed architecture, we have conducted experimental results on three benchmark datasets. Results demonstrate that our solution shows better performance with respect to existing approaches (e.g., a top–1 accuracy of 90.27% on the Food-101 challenging dataset).

Link of the paper: https://arxiv.org/pdf/1612.06543v1.pdf

**Methodology**

**Data:**

This dataset contains about 1000 images of pizza and 1000 images of dishes other than pizza. It can be used for a simple binary image classification task.

Number of instances in each class:  
Pizza: 983  
Not Pizza: 983.

All images were rescaled to have a maximum side length of 512 pixels. This is a subset of the Food-101 dataset > The Food-101 dataset consists of 101 food categories with 750 training and 250 test images per category, making a total of 101k images. The labels for the test images have been manually cleaned, while the training set contains some noise.

**Not Pizza CLASS**

**Not Pizza CLASS**

**Pizza CLASS**

**Pizza CLASS**

**Features:**

**• Mean:** Getting the mean intensity value of an image makes us know what color / value the picture intends to be, to get the intensity value of green channel image which help to identify the ground. Getting the mean is dividing the sum of the given intensity values of all pixels by the total number of pixels.

**• The median of intensity values:** Median value is the value in the middle of all given values. Ex: If the intensity values in the picture are 10, 15,20, 25 & 100, then the median is 20. This lets us know what value the mean tends to. It provides excellent noise reduction capabilities

**• Variance:** it removes all features whose variance doesn’t meet some threshold as it assumed that features as higher variance contains more useful information. Variance is a measurement of the spread between intensity values in an image

**• The standard deviation of intensity values:** Standard deviation measures how dispersed are the intensity values relative to the mean.

**• The skew of intensity values:** Skewness is how asymmetric is the normal distribution of the values. Each class may have similar skewness in most cases which simplifies classification

**• Extracting edge features:** This helps identify what objects are in the image. This helps identify the differences in the curvature of each class of objects.

**• Histogram equalization:** use it to improve the contrast of our images.

**• Canny edge:** is an edge detection operator that uses a multi-stage algorithm to detect a wide range of edges in images.

**• Thresholding**: way to select areas of interest of an image, while ignoring the parts we are not concerned with.

**Code:**

**Model:**

Support vector machines (SVMs) are a set of supervised learning methods used for classification, regression and outliers detection.

The advantages of support vector machines are:

• Effective in high dimensional spaces.

• Still effective in cases where the number of dimensions is greater than the number of samples.

• Uses a subset of training points in the decision function (called support vectors), so it is also memory efficient.

• Versatile: different Kernel functions can be specified for the decision function. Common kernels are provided, but it is also possible to specify custom kernels.

**Results:**

Accuracy :70% (8 features)

Accuracy :73% (6 features)

Accuracy :71% (2 features)

**Conclusion:**

In conclusion, we proposed a classification system that can separate different components of using the Machine learning tools. This system can be used to automatically classify pizza and not pizza. From the result, when tested against the pizza\_not\_pizza dataset, we got an accuracy of 73%. If more images are added to the dataset, the system accuracy can be improved in the future, we will tend to improve our system to be able to categories more items, by turning some of the parameters used.

**References:**

: https://arxiv.org/ftp/arxiv/papers/1606/1606.02210.pdf

: https://data.vision.ee.ethz.ch/cvl/datasets\_extra/food-101/static/bossard\_eccv14\_food-101.pdf

: https://arxiv.org/pdf/1612.06543v1.pdf