MEAL RECOGNITION

Team Member:

Kai Wu

Rundong Zhao

Ze Zhang

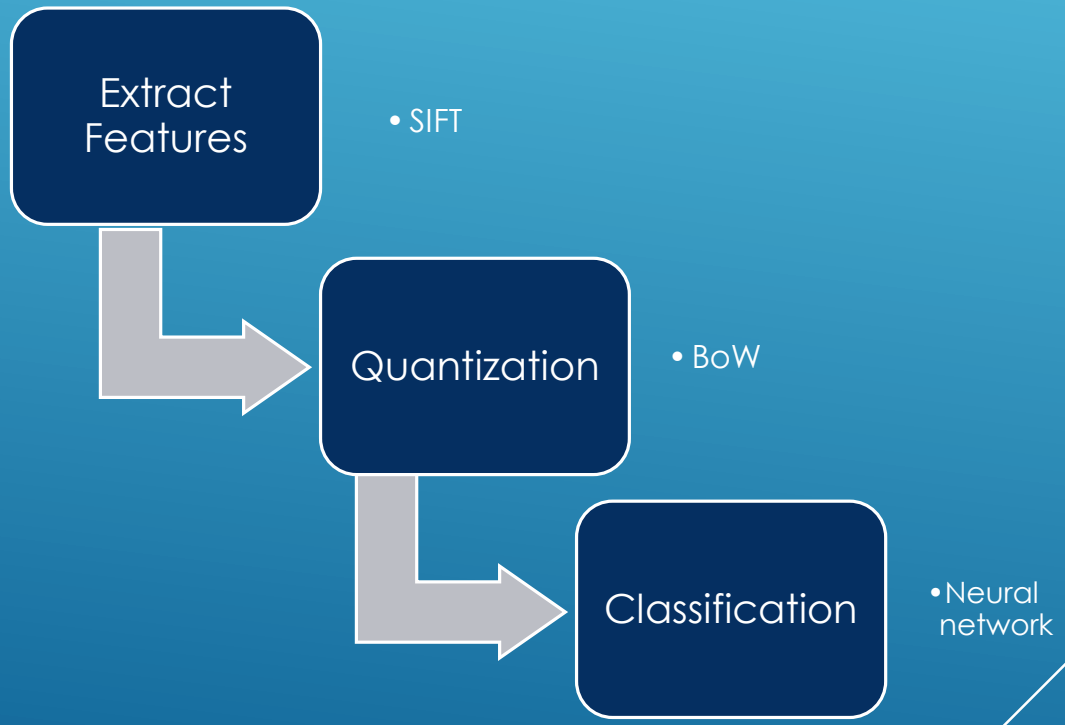
Introduction

Meal recognition technology becomes an important research topic because of its wide applications in other research fields, such as nutriology.

SIFT (Scale-Invariant Feature Transform), Bag of Words and Neural Network are used in our project.

Algorithm

System procedure



SIFT

SIFT feature is a 128 dimension vector, which we can view as a point in 128 dimension space.

For a certain image, after applying the SIFT algorithm, we can get several SIFT features for this image.

Note that different image may have different number of SIFT features.

Bag of Words

Collect all the SIFT features of all the images. Now we get many points in the 128 dimension space.

We do a k-means clustering to all the points, and get k centers.

k is specified by the user, and we select appropriate k value to make the result better.

Bag of Words

For each of the images, we calculate how many feature points belongs to each of the cluster centers.

We get a k dimensional vector for each of the image.

We normalize the vectors.

Now we have a normalized vector acting as a feature vector for each of the images.

Neural Network

Now we have the dataset

image 1: v_1 ,1

image 2: v_2, 1

....

image n: v_n, 10

The vector is the feature vector for the image, the index represents the meal type.

Use the neural network tool in matlab to train the data set.

Data Collection

We collected 10 classes of images from google image, about 60 training images per class.










Training Result

Apple	Banana	Broccoli	Burger	Fries	hotdog	pizza	Rice	Salad	Strawberry
0.79	0.86	0.95	0.74	0.87	0.74	0.78	0.87	0.90	0.88

Overall: 0.84

Test Result

Apple	Banana	Broccoli	Burger	Fries	hotdog	pizza	Rice	Salad	Strawberry
0.76	0.74	0.91	0.72	0.69	0.73	0.82	0.90	0.90	0.80

Overall: 0.80

Analysis

The performance of our classifier is fairly good. The required training time is approximately 4 hours. And the recognition speed is very fast, usually within 1 second.

SIFT does not consider the influence the background.