Hotel Image Classification Project Report

Author: Bhushan Suhas Deshpande

Contents

List of Figures

1 Introduction:

The objective of the final project was to be able to apply the knowledge of supervised learning to create a generalized classifier, which given an image will classify it in one of the following 8 classes:

1) Bathroom
2) Guestroom 3
3) Pool
4) Gym
5) Restaurant
6) Lobby
7) Aerial View
8) Business center

The training included 38372 images, each image belonging to one of the above 8 classes. Below is the example of the images of class “Pool”:

Figure 1: Sample Images

As it can be clearly scene the amount of pool visible in the 1st image and in the 2nd image is very different, makes the task of classification difficult.

2 Understanding the type of Classification:

There can be different types of image classification problem. E.g. object detection, scene detection, etc. So it was necessary to understand the classification problem. From the above mentioned example it is clear that the objects in the scene may be little different or may be in different number but the image all together can be well described as a scene. E.g. Below are the two images of restaurant.

Figure 2: Images for classification

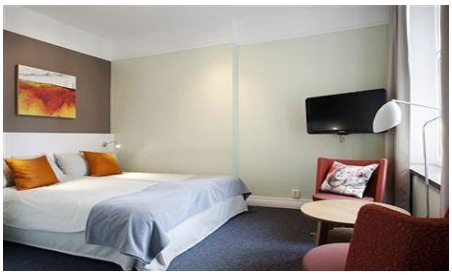
In the images, the number and shape of table and images is different. But their presence is enough to convince us that it is a restaurant scene. This helped me narrow down the techniques which can be used for image classification.

3 Approach:

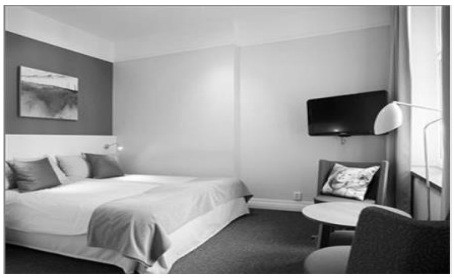
Bag of Features (also known as bag of words) is a very famous approach for scene classification in computer vision. In this approach the images are treated as word documents and image features as words. The histogram of these features represent an image. The approach is explained below:

Bathroom Image:

Bedroom Image:



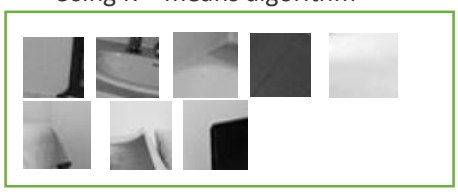
Turn it into grayscale



Feature Extraction using SURF

Generate Visual Vocabulary

Using K – means algorithm



Represent images as histogram of visual vocabulary

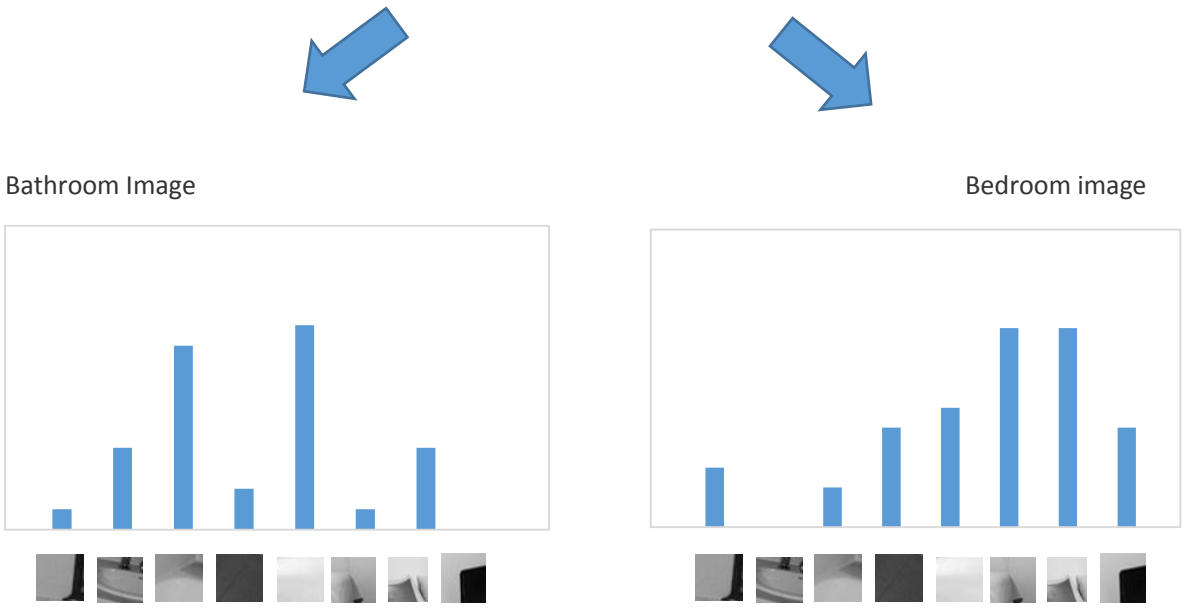


Figure 3: Image Representation approach using Bag of Features

4 Feature extraction:

The SURF algorithm has been used for feature extraction from images. The method is very sophisticated and it has following 3 parameters that can be tuned to get good results.

1. surf.upright = False

Setting it to false will make sure that a feature and rotated same feature will be considered as one and the same.

2. surf.hessianThreshold = 2000

One of the very important parameters, it controls the number of features extracted from every image.

Figure 4: Feature extraction plotted on image.

3. surf.extended = True

When set to true, features returned will be of size 128, otherwise it is of size 64.

5 Experiments PCA and K-means analysis (Visual vocabulary generation):

SURF returns features in the form of numerical values of 128 columns for each feature. In real, we do not need all the columns. Hence to speed up algorithm, I have done Principal Component Analysis and reduced the columns to 22. PCA was run on 10000 examples only.

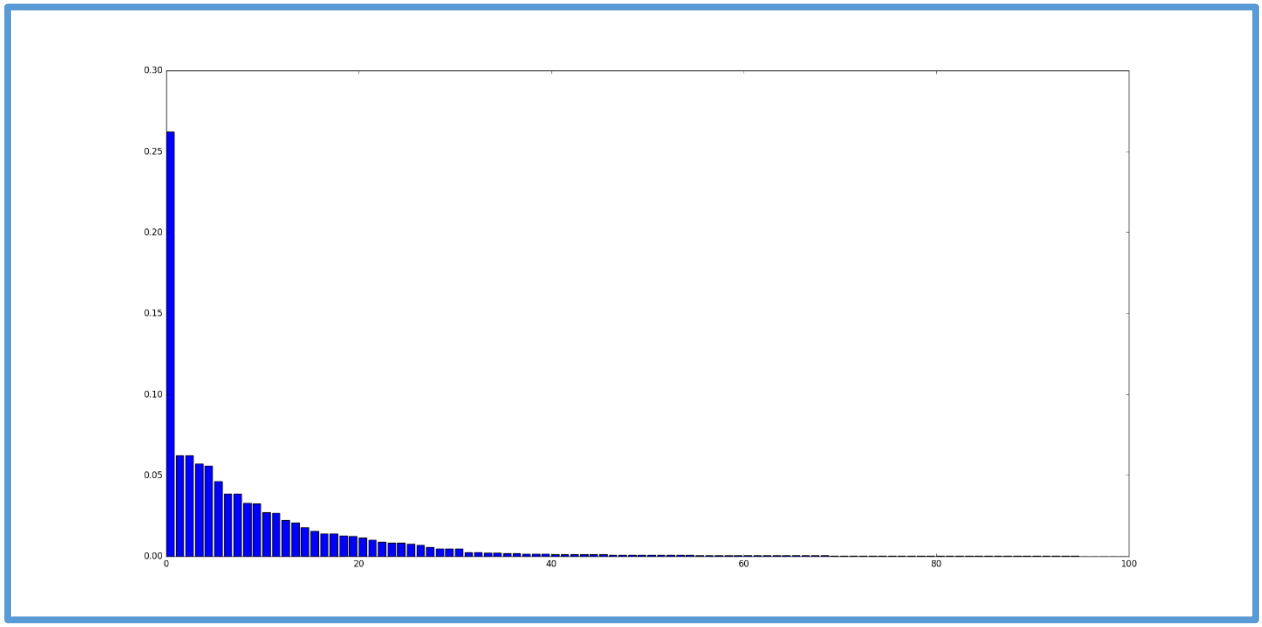


Figure 5: Principal Component Analysis: Variance vs components

K-Means is used to combine the features to come up with a consistent visual vocabulary. The histogram based on vocabulary is used to represent all the images as shown in figure 1.

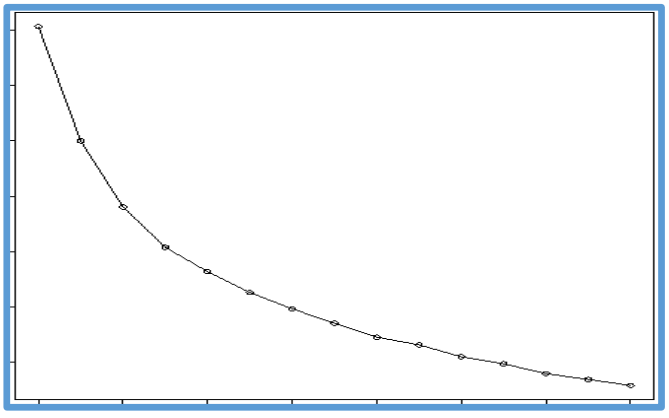


Figure 6: Trying to find elbow for k-means in a smooth curve.

Still according to my analysis, K=120 for the K-means from library skitlearn could have been good. But due to bad time complexity of code, I had to settle with k=15, k=30 only. Also the k-means was run repeatedly, so that it does not get stuck in local minima.

6 Results:

6.1 Random Forest for Classification:

Random forest is a descent classifier for image classification. It was tuned for 300 decision trees and default length. The number of trees were restricted to 300 to avoid overfitting on training data. Random forest become immediate favorite as it requires less tuning and gives a descent performance.

The other classified tried was SVM with kernel as Radial Basis Function. Because of probably bad tuning, it did not do better than random forest for classification.

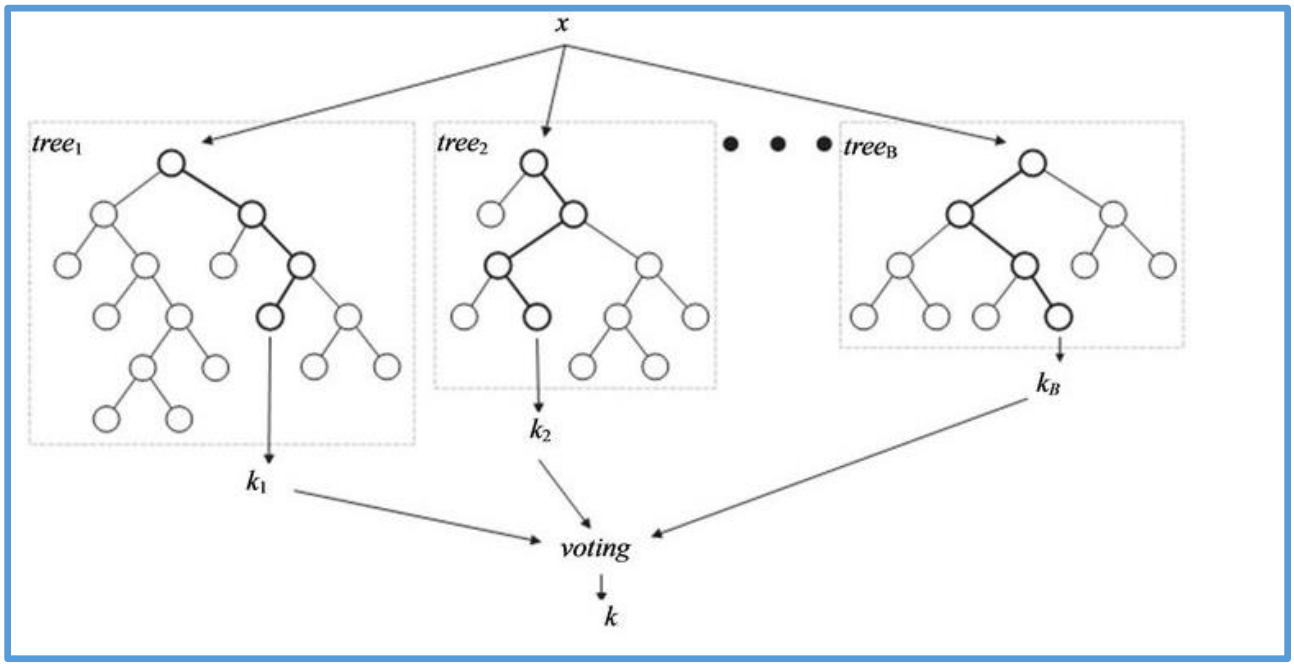


Figure 7: Random forest algorithm

6.2 Accuracy:

The validation showed the accuracy of 55%. The same is shown in the Kaggle competition with the score of 1.46

7 References:

The video for Bag of feature explanation.

https://www.youtube.com/watch?v=iGZpJZhqEME