Scene categorization using Naive Bayes Classifier

Mahmut DEMIR

CMPE 547 Computer Engineering Department - BOGAZICI UNIVERSITY

Problem

Scene categorization is a fundamental problem in computer vision. Most approaches require human experts to annotate the training images. In our application, we aim to decide scene category of given image in an automated manner.

A scene can be described in terms of a set of objects (e.g. car, window, tree, sky) that can be found inside and in terms of their quantities. After enough training with representative images from various scene categories, we can obtain scene-object joint probabilities table. In such table, particular set of objects will have a higher probability of existence (equivalently absence) given a particular scene category.

SUN image database provides a comprehensive collection of annotated images that covers a large variety of environmental scenes, places and the objects within. Referring SUN database, we created a table which contains 50 scene categories and ratio of occurrences of 25 common objects for each category as shown in table 1.

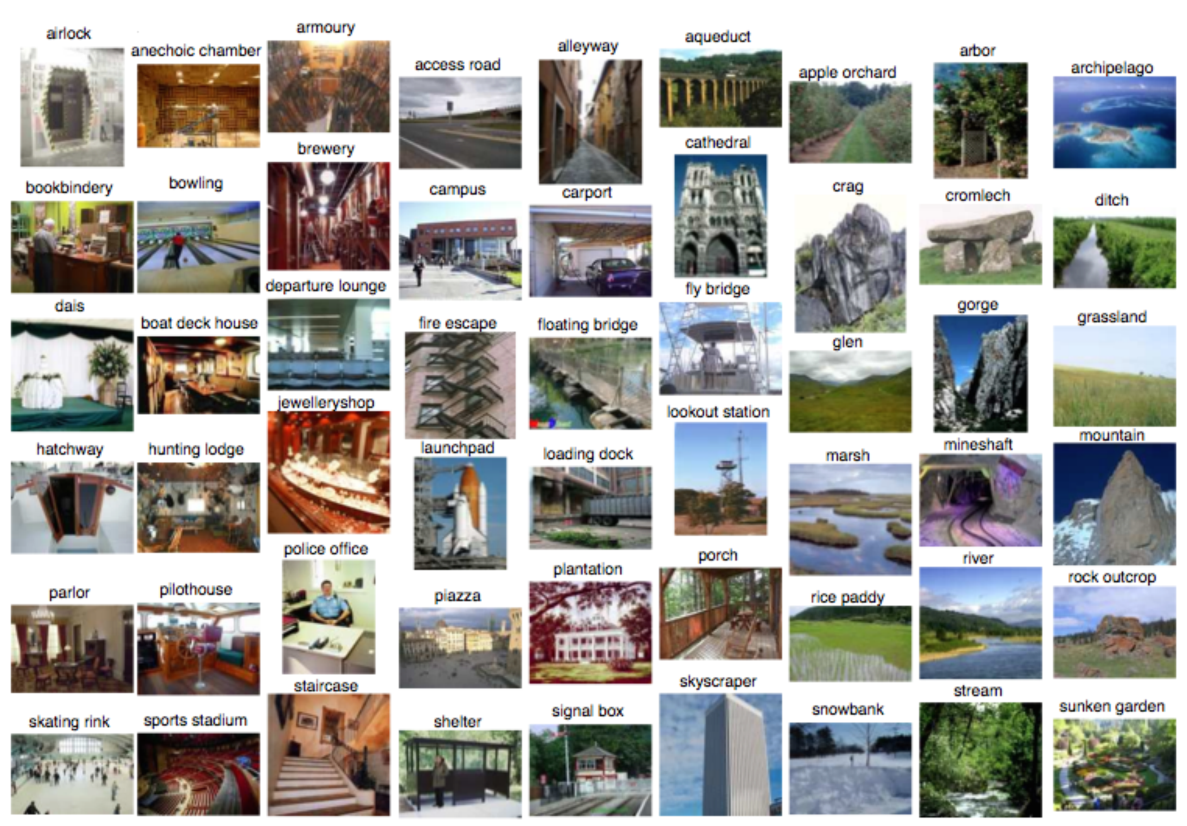

Figure 1: Example of scene categories

We use Naive Bayes Classifier (NBC) in order to classify given images based on objects detected within. Strong independence between objects given a particular scene category is assumed and inference is carried out using maximum a posteriori decision rule.

In order to test our approach, test user is provided with a random image and asked to choose a number objects that exist and do not exist in current image. Depending on evidence prompted from user, program outputs five best-matching scene categories.

Methods

1. Probabilistic Model

Test user is provided with a random image and prompted to list some objects exist and some objects do not exist within image. Each input will be used as evidence for our probabilistic model and because we use NBC we assume independence between objects conditioned on scene category.

$$p(C_k|e_1, e_2, ..., e_n) \;\; e_i \in \{o_j, \tilde{o_j}\} \tag{1}$$

$$where\ i = 1...N_{evidence},\;\; j = 1...N_{object},\;\; k = 1...N_{scene}$$

$$p(e_i = o_j|C_k) \Rightarrow Probability\ of\ j^{th}\ object\ \underline{exist}\ in\ scene\ category\ k$$
$$p(e_i = \tilde{o_j}|C_k) \Rightarrow Probability\ of\ j^{th}\ object\ \underline{do\ not\ exist}\ in\ scene\ category\ k$$

Conditional probabilities are calculated as follows:

$$\begin{aligned} p(e_i = o_j|C_k) &= \frac{N_{object_j k}}{N_{scene_k}} \\ &= \frac{total\ \#\ of\ object\ type\ j\ present\ in\ images\ of\ scene\ category\ k}{total\ \#\ of\ images\ of\ scene\ category\ k} \end{aligned} \tag{2}$$

By using “naive” conditional independence assumption of naive bayes, we can express joint probability as:

$$\begin{aligned} p(C_k|e_1, e_2, ..., e_n) &\propto p(C_k, e_1, e_2, ..., e_n) \\ &\propto p(C_k)p(e_1|C_k)p(e_2|C_k)...p(e_n|C_k) \\ &\propto p(C_k)\prod_i^n p(e_i|C_k) \end{aligned} \tag{3}$$

$$p(C_k|e_1, e_2, ..., e_n) = \frac{1}{Z}p(C_k)\prod_i^n p(e_i|C_k) \tag{4}$$

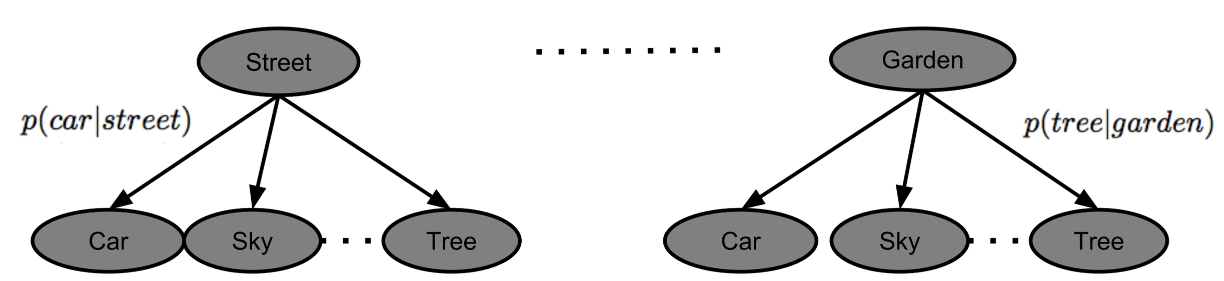

Figure 2: p(object|scene) conditional probabilities are calculated using Table 1

2. Inference

Posterior probabilities for each scene categories calculated based on a number of evidences given naive bayes assumption. We then use maximum a posteriori approximation for finding most probable scene category.

$$C_k = \underset{k \in \{1...K\}}{\arg\max}\, p(C_k) \prod_i^{N_{evidence}} p(e_i|C_k) \tag{5}$$

Results

Android mobile application was used for testing our algorithm. For this work, data table of 50 scenes and 25 objects shown in table 1 embedded in the software and conditional probabilities were inferred from the table using eq.3. Sample images which represent various scene categories obtained from SUN image database and embedded in software.

		# of images	Wall	Window	Chair	Floor	Sky
Outdoor Natural	Picnic Area	1		1			1
	Playground	9		4			2
	Beach	118					111
	Forest	127		11			76
Transportation	Backstairs	10		8	2		4
	Bicyle Roads	11		6			
	Bridge	24					18
	Bus Station	7		22			6
	Gas Station	20	7	25			20

Table 1: Part of table of 50 scenes and 25 object occurrences

For the test scenario, user was prompted a scene image which was randomly selected from 20 sample images. First, user was requested to choose a number of objects out of 25 which exist in the current image. Next, user selected a number of objects which do not exist in the image. Five best matching scene categories with associated probabilities were calculated and shown on the screen after each evidence entry.

Scene	1 object	2 objects	3 objects	1 object	2 objects	3 objects
Bazaar Outdoor	Y(7.0)	N	N	N	N	N*(9.0)
Museum Outdoor	N(7.0)	N*(10.7)	Y(14.0)	Y(19.0)	Y(22.3)	N*(28.8)
Construction Site	N	N	N	N	N	N
Botanical Garden	Y(5.0)	Y(9.5)	Y(10.76)	Y(11.0)	Y(15.0)	Y(17.0)
Apartment Building	N	N	N	N	N	N
Building Facade	N	Y(6.8)	Y(7.0)	N	N	N
City	N*(5.0)	N*(6.5)	N*(13.39)	Y(21.0)	Y(39.9)	Y(41.0)
Downtown	N*(5.0)	N*(7.0)	N*(14.0)	Y(19.0)	Y(20.0)	Y(22.8)
Gas Station	Y(5.0)	Y(7.0)	Y(8.4)	N*(10.7)	N	N*(10.9)
Motel	N	N	N*(7.8)	N*(10.2)	N*(10.4)	N*(11.0)
Street	N	N	Y(16.0)	Y(36.0)	Y(44.0)	Y(62.0)
Schoolhouse	N	N*(7.0)	Y(9.6)	Y(11.2)	Y(20.6)	Y(20.7)
Convenience Store	Y(5.0)	Y(7.0)	Y(13.0)	Y(20.0)	Y(24.0)	Y(27.0)
Swimming Pool	N	Y(9.0)	Y(15.0)	Y(16.0)	Y(23.0)	Y(30.0)
Total Avg.	4/2.78	5/5.03	8/9.21	8/12.43	8/15.65	7/20.01

Table 2: Scene category recognition results.
Y(x) = Category recognized successfully with probability x. N*(x) = Category not recognized but among five best matching. Last row shows recognition ratio along with probabilities of success.

It was clearly observed that each additional evidence entry “generally” increases success rate in terms of increasing probability of true scene category in the estimation. There are several important points that affect the success rate of estimation:

1) Training data is not always enough! Sample image rarely contains objects not representative of scene category image belongs to.
2) Some objects(e.g. window, tree) have multiple occurrences in some scenes.
3) Some scene categories may contain similar objects with similar conditional probabilities.

Conclusions

In this work, we applied Naive Bayes Classifier to classify a given image into one of predefined scene categories. We used an open source image database which contains large amount of annotated images with over 900 scene categories. For our case, we applied categorization for 50 scenes. Classification was based on objects detected within image and class conditional probabilities were extracted from scene-object relation table we created. In order to test our algorithm, we developed an android mobile application in which test user was shown a random image and requested to enter some objects exists and do not exist within given image. After each evidence entry, best matching scene categories shown to user.

As we stated in previous section, success rate of our algorithm was directly related to training data. In some of case, user is shown an image of some category but some objects within image might not be in the training data of that category. Second, some scene categories have identical object occurrence distribution with other categories which causes category of given image not recognized correctly. Third, some objects such as windows or trees in an image has multiple occurrences in images. Because these occurrences increase conditional probabilities, stating object existence or absence might not be enough. Instead, it would be wise to state number of occurrence of object either qualitatively(small, many) or quantitatively(1,2,3..).

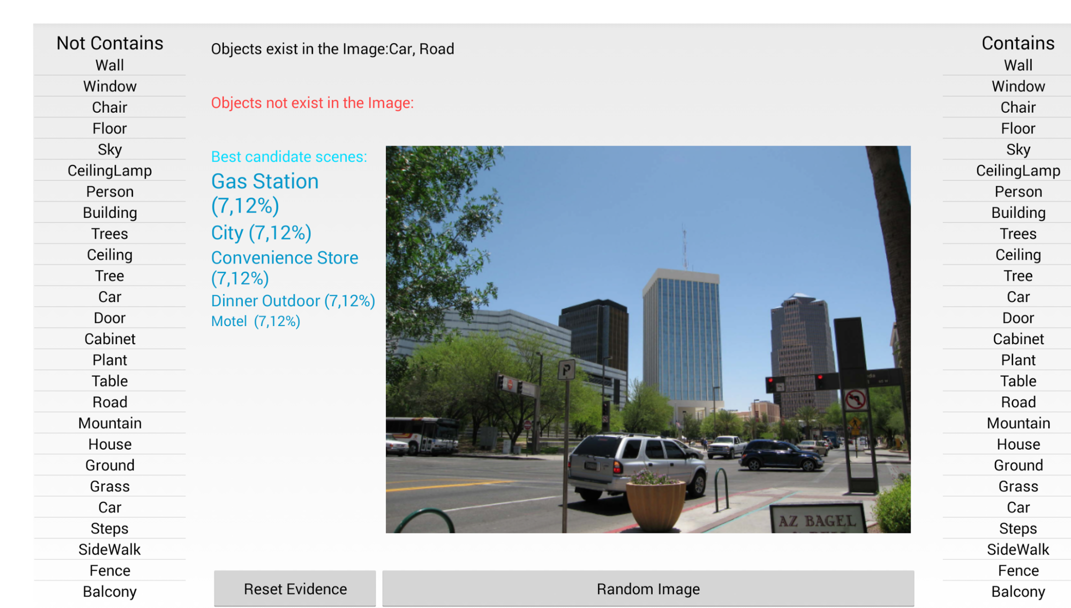

Figure 3: Android mobile application screen

Bibliography

1. "SUN Database: Large-scale Scene Recognition from Abbey to Zoo". J. Xiao, J. Hays, K. Ehinger, A. Oliva, and A. Torralba. IEEE Conference on Computer Vision and Pattern Recognition, 2010.
2. "Cognitive maps for mobile robots—an object based approach." Vasudevan, Shrihari, et al. Robotics and Autonomous Systems 55.5 (2007): 359-371.
3. “Naive Bayes Classifier” http://en.wikipedia.org/wiki/Naive_Bayes_classifier