Color Classification and Recycling Bin Detection

Parthasarathi Kumar
Dept. of Electrical and Computer Engineering
University of California San Diego
San Diego, USA
pakumar@ucsd.edu

I. Introduction

Object detection is an important problem in robotics and is used extensively for different tasks such as localisation, collision avoidance, tracking, path-planning, etc. It involves locating object(s) of interest and accurately estimating its position in the image. There are various approaches for this task. In this report we use probabilistic models trained in the color feature space to detect blue recycle bins in images. More specifically we first take up the task of classifying single color images as red, blue or green. We then extend this approach to classify image pixels as blue recycle bin or not which is in turn used to detect these blue recycle bins of interest.

II. Problem Formulation

A. Color classification

The problem of color classification takes in a nx3 RGB triplets as input and outputs a class label y = $\{1, 2, 3\}$ for each of the input triplet where output labels 1, 2 and 3 correspond to red, blue and green color respectively.

B. Recycle bin detection

The problem of blue recycle bin detection takes in a RGB image as input and outputs a list of bounding boxes. For each image, the output list can be of size 0 or more i.e. an image may contain no blue recycle bins or 1 or more of them. For each recycle bin the output bounding box is of the form $[x_1, y_1, x_2, y_2]$ where (x_1, y_1) and (x_2, y_2) represent the top left and bottom right coordinate respectively

III. Technical Approach

A. Color classification

For color classification we train 3 multivariate Gaussian distributions, 1 for each class. We use the Maximum Likelihood Estimate to model the distribution leading to the following equations -

$$\mu^* = \frac{\sum_1^N x_i}{N} \tag{1}$$

$$\Sigma^* = \frac{\sum_1^N (x_i - \mu)^T (x_i - \mu)}{N} \tag{2}$$

It is evident from (1) and (2) that the mean and co-variance of the Gaussian distribution are the sample mean and co-variance respectively. The prior probabilities of each class is the relative frequency of the respective class in the training data. To train the model, we use the provided data-set which includes 1000+ images across each class, where each image consists of variants of a color class.

For testing/inference, we calculate the posterior probabilities of each class and assign the class label corresponding to maximum posterior probability. More formally,

$$y^* = argmax_y P(y|x; \mu^*, \Sigma^*) \tag{3}$$

where y = 1,2,3.

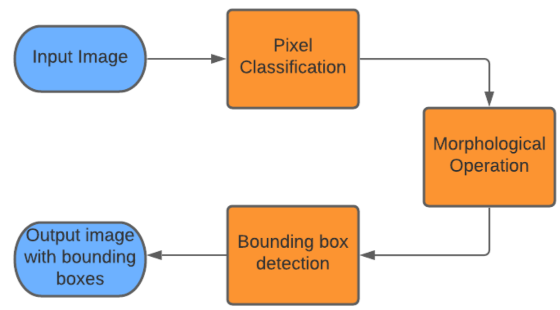

Fig. 1. Pipeline for the bin detection

B. Recycle bin detection

For recycle bin detection, we employ a color based segmentation algorithm followed by shape detection to detect the recycle bins as described in Figure 1.

The segmentation module is an extension of the pixel classification approach where we train 2 multivariate Gaussian distributions on bin-blue non bin-blue examples.

1) Training data: Training data for this task is collected from the given images. We mark the 2 classes on each image, and a average pixel value is calculated for each marked region, which is used for training purposes. Given the variation in lighting conditions, we train our model in the HSV colorspace which separates the intensity information from the color information unlike RGB colorspace. We collect approximately 158 negative samples and 58 positive samples to train our model. The ratio of 3:1 (approx.) data points in the negative and positive classes is a fair estimate of the 2 classes which

TABLE I
PARAMETERS OF GDA FOR PIXEL CLASSIFICATION

Red	Green	Blue
$\mu = [0.75 \quad 0.34 \quad 0.34]^T$	$\mu = [0.35 \quad 0.73 \quad 0.32]^T$	$\mu = [0.34 \quad 0.33 \quad 0.73]^T$
$\Sigma = \begin{bmatrix} 0.037 & 0.018 & 0.018 \\ 0.018 & 0.0619 & 0.0085 \\ 0.018 & 0.085 & 0.0620 \end{bmatrix}$	$\Sigma = \begin{bmatrix} 0.055 & 0.017 & 0.008 \\ 0.017 & 0.034 & 0.017 \\ 0.008 & 0.017 & 0.056 \end{bmatrix}$	$\Sigma = \begin{bmatrix} 0.054 & 0.008 & 0.017 \\ 0.008 & 0.056 & 0.018 \\ 0.017 & 0.018 & 0.035 \end{bmatrix}$

TABLE II
PARAMETERS OF GDA FOR BIN CLASSIFICATION

Bin-Blue	Non Bin-Blue
$\mu = [0.26 \quad 0.32 \quad 0.56]^T$	$\mu = [0.42 \quad 0.74 \quad 0.66]^T$
$\Sigma = \begin{bmatrix} 0.025 & 0.026 & -0.002 \\ 0.002 & 0.057 & 0.0016 \\ -0.000 & 0.001 & 0.006 \end{bmatrix}$	$\Sigma = \begin{bmatrix} 0.002 & 0.003 & 0.001 \\ 0.003 & 0.037 & -0.008 \\ 0.001 & -0.008 & 0.034 \end{bmatrix}$

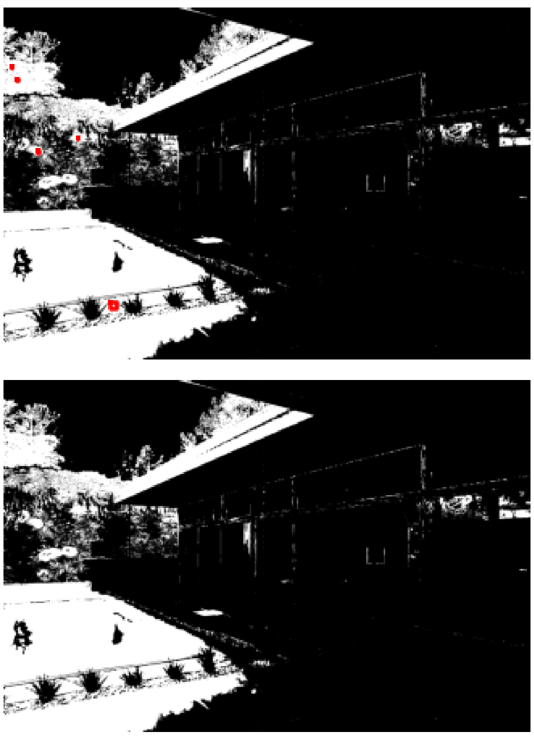

Fig. 2. Reduction of false detections on suppressing candidates based on relative area.

helps us in achieving decent results as discussed in Section IV.

2) Segmentation Module: The segmentation module, as indicated earlier is a binary classifier based on multivariate Gaussian Distribution trained on the data collected for this task. We use a vectorized implementation to predict the class of each pixel in approx 0.3 seconds for each image.

3) Bin detection: Once the binary mask is obtained, we apply an erosion operation using a square kernel to refine the mask. This leads to incorrectly marked blobs to filter out resulting in removal of a large number of false detections. We parameterize the kernel size with respect to the image size in order to make it effective across image dimensions.

Following this we use connected component labeling to label each region in the mask which in turn is fed into region props function to calculate properties of each candidate region. We utilize the metric of aspect ratio to eliminate any existing false positives and localize out object of interest. The main idea behind using aspect ratio is to exploit the prior information of standard recycle bins being longer in shape i.e. height $>$ width or depth. We further suppress some candidates if they are too small.

IV. RESULTS

A. Pixel classification

For pixel classification task, we achieve a perfect score of 1 on the validation set and a score of 9.934/10 on test set. Refer Table 1 for the parameters for each of the model.

B. Recycle bin detection

For recycle bin detection we achieve a score of 9/10 on the validation set and a score of 8.5/10. Refer Table 2 for parameters for the model.

Although the segmented mask is decently accurate, it still has a large number of false detections (i.e. marked as bin-blue). To eliminate these, we apply a erosion (morphological) operation as indicated in Section III. This leads to good improvement, but still leads to a significant number of incorrect detections. A possible reason for this was the hard-coding of the kernel size that was being used to erode the image. To overcome this we use a image dependent kernel size to achieve better results

Fig. 3. Elimination on false detections based on aspect ratio of candidate bouding boxes.

In addition to this, another metric of area was applied to eliminate very small bounding boxes as shown in Figure 2. To be more specific, any candidate with a area (in $pixel^2$) $<$ 0.1% of the image area is suppressed.

Figure 3 shows the elimination of candidates that do not satisfy the condition of aspect ratio

Figure 4 & 5 depicts the results on the validation set. First column indicates the images, followed by the binary mask obtained from pixel classification and erosion operation. The last column shows the images with the detected bounding boxes.

REFERENCES

[1] ECE 276 A Lecture Slides.
[2] Scikit Image Documentation
[3] OpenCV Documentation

TABLE III
BOUNDING BOXES ON VALIDATION SET

Validation Image Id	Bounding boxes
61	$\begin{bmatrix}125 & 185 & 286 & 309\end{bmatrix}$
62	$\begin{bmatrix}188 & 232 & 438 & 413\\ 349 & 27 & 499 & 132\end{bmatrix}$
63	$\begin{bmatrix}97 & 175 & 225 & 265\end{bmatrix}$
64	$\begin{bmatrix}109 & 348 & 272 & 462\end{bmatrix}$
65	$\begin{bmatrix}425 & 829 & 615 & 914\end{bmatrix}$
66	$\begin{bmatrix}. & . & . & .\end{bmatrix}$
67	$\begin{bmatrix}315 & 721 & 500 & 815\\ 317 & 599 & 497 & 687\end{bmatrix}$
68	$\begin{bmatrix}. & . & . & .\end{bmatrix}$
69	$\begin{bmatrix}. & . & . & .\end{bmatrix}$
70	$\begin{bmatrix}. & . & . & .\end{bmatrix}$

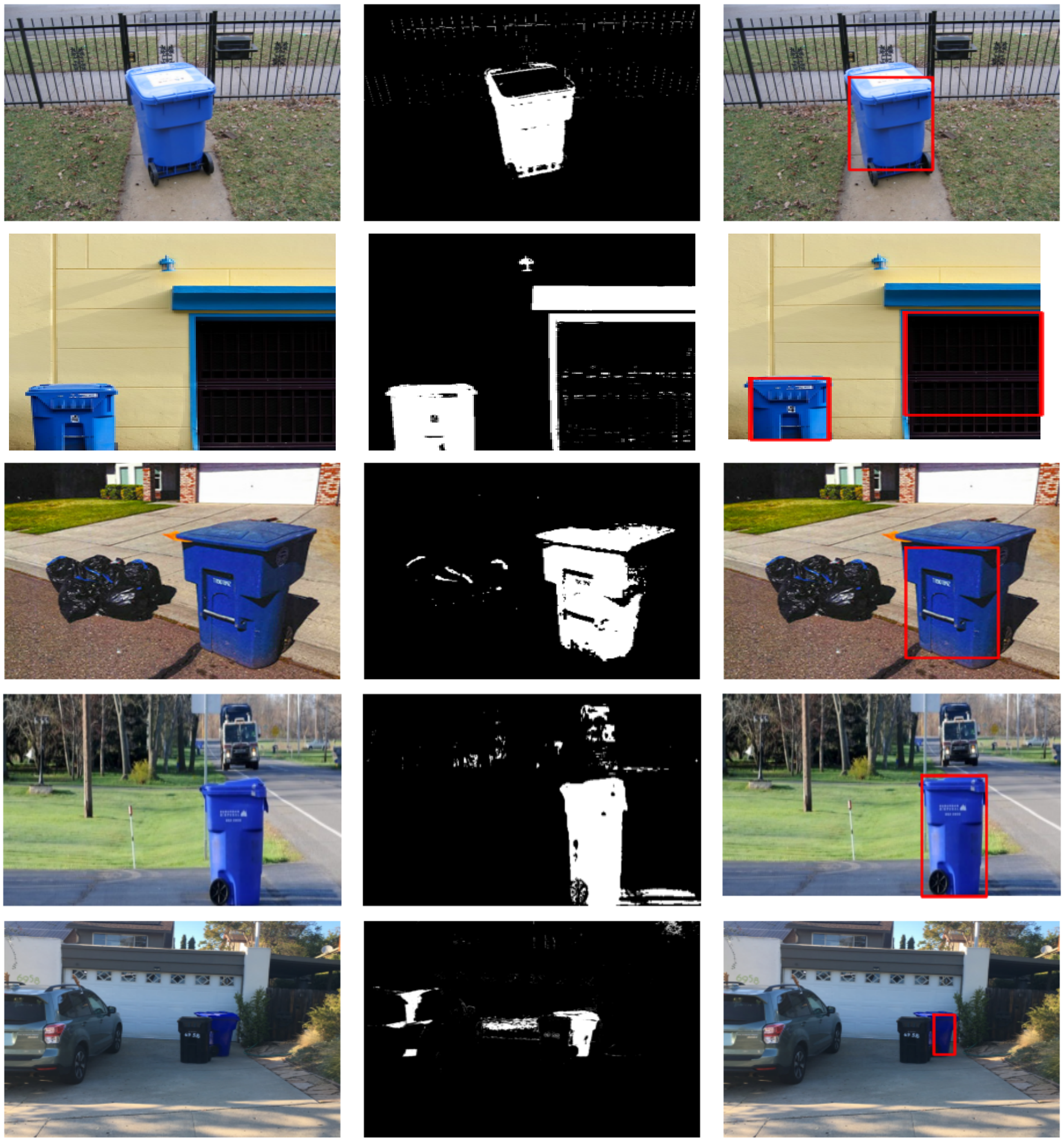

Fig. 4. Results from validation set. Coloumn 1 shows the original image, coloumn 2 indicates the mask and the column 3 shows the detected bouding boxes

Fig. 5. Results from validation set. Coloumn 1 shows the original image, coloumn 2 indicates the mask and the column 3 shows the detected bouding boxes