

# Garbage Classifier

## EECS 4422 Computer Vision

### Scientific Stream Project

Author: Koko Nanahji  
Student Number: 215168057

#### Abstract

This paper describes the implementation of a garbage classification tool that can be used to sort garbage in a three-bin garbage segregation system. This tool is designed to be used as part of augmenting the existing garbage bins into “smart” garbage bins that could replace the bins that are located on our streets. The algorithm uses a classical computer vision method called Histogram of oriented gradients (HOG) to represent each image and uses Support Vector Machine as the classification algorithm. The data used in this project consists of images that were found on online repositories and other images that were specifically captured for this project.

#### 1) Motivation

Environmental protection has become one of the major concerns of 21<sup>st</sup> century. One of the main contributors to the pollution and contamination of the environment is insufficiency in waste management systems. Implementing an effective waste management infrastructure can help us clean up our environment and preserve it for future generations. Successful waste management starts from the time that the object makes it to disposal. Therefore, separating waste into proper categories is a crucial part in the overall waste management infrastructure. However, finding the right garbage bin for a specific object can be challenging and learning where each object should be placed can be time consuming. The main motivation behind the project described in this paper is to help ease this process. This garbage classifier application is designed to help categorize disposed items in a three-bin garbage segregation system.

#### 2) Context

The application described in this paper is designed to be part of the augmentations to create “smart” garbage bins. In order to benefit from this application, a new set of hardware devices need to be added to the current garbage bins that are located on the streets. The main job of the added hardware is to capture an image of the disposed object and feed the image to this application. Afterwards, the application classifies the object and returns the ID

corresponding to a garbage bin. The added hardware, after receiving the ID of the identified bin, drops the object into the correct bin. Therefore, upgrading current garbage bins into “smarter” ones should ease the process of garbage segregation and help us better take care of our environment.

### 3) Dataset Description

This application is developed for the three-bin system, where there is one bin for recyclables, another bin for organics, and one additional bin for non-recyclable items. Therefore, the dataset to train and test the application contains items that belong to each category. Given time restrictions for this project we chose only a few items that are typically seen in garbage bins. This project’s dataset contains apple and banana images which belong to the organic category. It also contains images of cans, such as soda cans, tomato soup cans and chowder cans as well as juice boxes (both small and large boxes) which all belong to the recyclable category. In addition, it contains images of (standard) light bulbs in the non-recyclable category. This dataset consists of images that can be found on computer vision repositories (links can be found in the ReadMe.md file) and additional images that were taken manually to increase the variety of the items in the dataset. One of the major issues with gathering this dataset is the disproportionate number of images that exists for certain categories compared to the others. For example, there are many apple and banana images on computer vision repositories. However, images of light bulbs are almost non-existent. Therefore, image augmentation techniques were applied to the categories that suffer from limited dataset size. These augmentations were simple affine transformations, such as rotation and scale changes, which helped in creating a better dataset.

### 4) Method

The garbage classification algorithm described in this paper is based on classical computer vision and machine learning methods. The main reason for using classical methods instead of using newer and more popular techniques such as deep neural networks is the limited size of the available dataset. Since neural networks are “data hungry”, training a neural network would require a large volume of training data. However, classical methods are proven to achieve “reasonable” success rates without requiring as much data.

The classification method that is used in this application is Support Vector Machine (SVM). The main idea behind Support Vector Machines is to find the hyperplane that has the largest distance to the nearest training data point of any class in the input space. Since, in general it is believed that the larger the margin (the distance from the hyperplane to the nearest training data point of any class), the lower the error of the classifier will be in real world data. The reason that Support Vector machines were chosen as the classification method is because the evaluation results showed that, for this specific problem and dataset,

Support Vector Machine performed slightly better than the other two algorithms that were originally being considered as the classification method (K-Nearest Neighbours algorithm and Multi-class logistic regression).

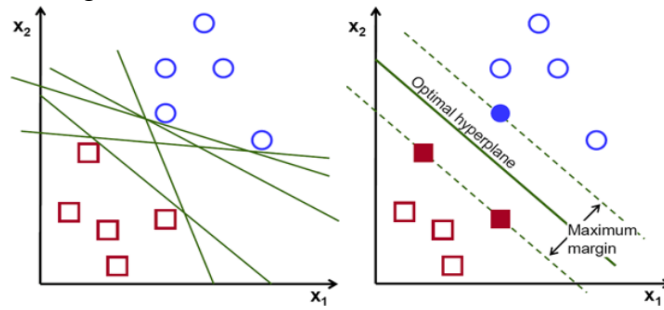


Figure 1: Shows the hyperplane chosen by Support Vector Machine in this simple example

Support Vector Machine takes a feature vector describing each image as its input. This feature vector is created by the Histogram of Oriented Gradients (HOG) method. The main idea behind this method is to divide the image into small patches and create a histogram of the oriented gradients for each patch. The combination of these histograms constitutes the feature vector that is used to describe each image. The reason that the Histogram of Oriented Gradients algorithm is used to create the feature vector is because it outperformed the other technique that was tried, which is matching keypoints technique. However, given the shapes of the items, the keypoint approach performed poorly and did not scale properly even on a small set of items.

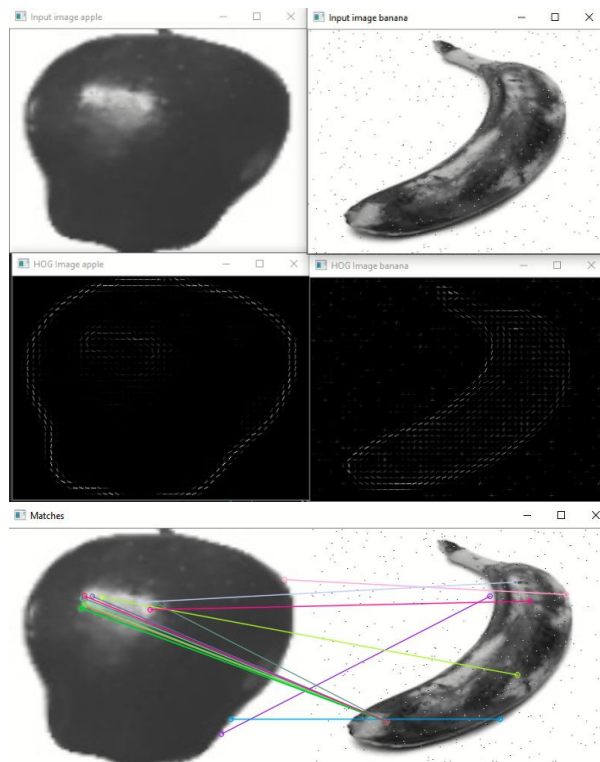


Figure 2: Example to visualise the difference between HOG and Keypoint descriptor. First row shows the input images. The second row shows the output of HOG on corresponding input image. The third row shows the selected keypoints in each input image and links the keypoints that were matched by Flann (in cv2)

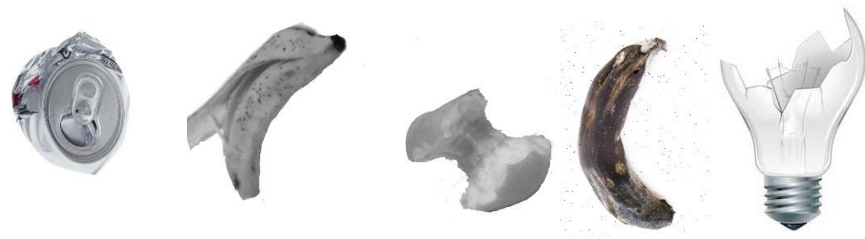
Additionally, given the disproportionate size of data for each category in the dataset (even after data augmentation), we were only able to use small subset of the images in the categories that had excessively more data than the other categories. For example, as was mentioned before, there are a lot of (about 1000) images for apples. Due to lack of images in the other categories, such as light bulbs, we only used small sample of (first 150) images from the apple category for training the model to avoid the problems that would be caused by the disproportionate input sizes.

## 5) Evaluation Method

The evaluation of the model is based on k-folds cross validation method. Cross validation is used in measuring the prediction accuracy of a predictive model and estimating the performance of the model in real world. The k-fold cross validation divides the original dataset into k equal sized chunks. One of the k chunks is used for testing the model, and the remaining  $k - 1$  chunks are used as the training data. This process is repeated k times, with each of the k subsamples used exactly once as the test data. Therefore, the combination of the k results can be used as a single estimation of the accuracy of the model. The main advantage of this method over other methods is that all the available data is used for both training and testing. Therefore, given the limited size of the dataset, cross validation was a reasonable choice for model evaluation for this project.

## 6) Experimental Results

The dataset contains images of various categories to appropriately evaluate the proposed method. These categories consist of apple, banana, small juice box, large juice box, light bulb, chowder can, soda can, and tomato can. Additionally, the dataset contains images of rotten apple, rotten banana as well as crushed soda cans and light bulbs. The other cans are much more solid and harder to crush therefore they are not included in the dataset (could be added in the future). It is noteworthy to mention that the original project proposal contained garbage bags as one of the other categories of input. However, due to the time restriction as well as the lack of data this item was removed.



*Figure 3: Shows example of images in the dataset.*

## 7) Analysis of Results

The results gathered from cross validation shows that the proposed method (i.e. HOG with SVM) has reasonable success rate (about 97%) in choosing the correct garbage bin for given item. Moreover, it provides empirical evidence that this approach might be used as an alternative for neural networks in projects that suffer from lack of training data. Further analysis on the misclassified object shows that the objects that have the same shape tend to be misclassified. The main reason for this limitation is that the approach taken in this paper is based on HOG descriptors and these descriptors only consider the shape of the object. Therefore, the objects that have similar shapes will potentially be classified incorrectly. For example, banana and light bulb are the most incorrectly classified objects and the dataset contains images of these items that have very similar shapes (As can be seen in Figure 4).

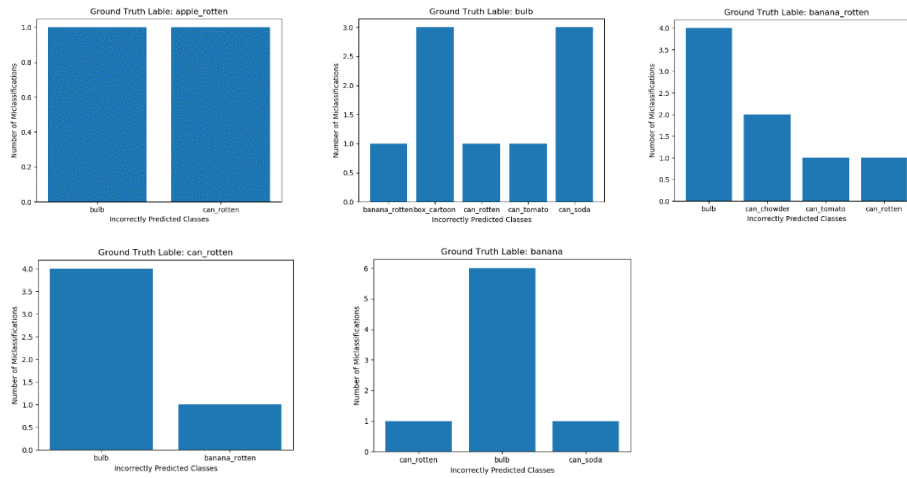


Figure 4: Shows bar plot of misclassification of the proposed computer vision method.

This issue will also limit the ability to scale the proposed method to larger set of items. The main reason is that new items that have similar shape as the existing items will cause the algorithm to have poor performance. One approach to overcome this issue could be to add other features, such as color and texture, to the feature vector of each image.

## 8) Restrictions and Future Work

One major restriction of the approach described in this paper is that it can have difficulty classifying objects that look identical (not just similar shapes) but belong to different categories. For example, real apple and plastic apple look exactly alike (which is main goal of the plastic apple producers), however real apple belongs to the organics category and plastic apple belongs to the recyclable category. Therefore, using the image alone to distinguish these objects might not provide enough information to classify them correctly. In order to rectify this problem, we could add another metric into the feature vector of each item, such as the weight of the object, and make the predictions accordingly.