Masa Hu
May 5, 2017
CSCE 470

Self Reflection

For the Trash Classification Project my main role was designing and implementing the image processing and classification software. Implementation of the software required the use several open source libraries and resources. These included svm_light, python-opencv, tkinter, matplotlib, and python-numpy. Most of the software was written in python but in order to stitch the usage of svm_light with the written python programs, bash scripts were also implemented to simplify the steps needed to use the software. I would say creation of the classification software and producing acceptable results was 60% of the project and by the end of the project I had finished 2/3 of the work assigned to me.

A big obstacle for this project was the creation of a design for trash classification that could produce accurate results because image classification is still a difficult problem that has not been solved yet. Fortunately, having some background in computer vision I immediately knew to use the opencv library for image processing. I implemented a support vector machine as the machine learning algorithm for classification because I knew how to implement svm_light from a previous project and I wanted to spend as much time as possible on producing results for this project instead of exploring different algorithms. The solution that my group and I came up with using the advice of our sponsor was to use color distribution of segments of the images of trash and see if those could be used for training and classification.

Looking at the assessment rubric, I've improved the most on Outcome 3, 8, and 11. One of the innovative approaches I decided to use to create the training data was to use images completely filled with a certain classification category. I recognized the necessity to have preliminary results based on the current design of the software since image classification is not a solved problem. As the project progressed I discovered many more tools in the opencv computer vision library, and I had to upgrade my version of opencv to the latest in order to use a segmentation function I needed. I recognize that software will continue to develop and continuing practice with new tools is necessary to keep up. During the project I made it an effort to make my code as reusable as possible in consideration with the future work that may happen with this project.

The hardest to criteria to overcome was Outcome 2, 9, and 10. Identifying the problem is difficult when the problem is not yet solved. I did not have much time to learn a different machine learning algorithm and instead went with a support vector machine. At the end of the day, I wanted the program to function properly more than having it use the latest tools and techniques. I would try to keep my code base as efficient as possible while moving along, but I'm certain I missed some things.

Although I was able to implement the classification software, I was not able to produce the accuracies desired for the project. It is my current thought that the features used to describe the images, namely HSV color distribution, may not have enough information about the images in order to distinguish the different categories from each other. It may also be the fact that the usage of a different machine learning algorithm besides SVM may improve accuracy results as well. However, I did not have enough time to change the feature extraction process or implement a different machine learning algorithm while improving the code base and producing results.