Image Recognition of Trash Objects

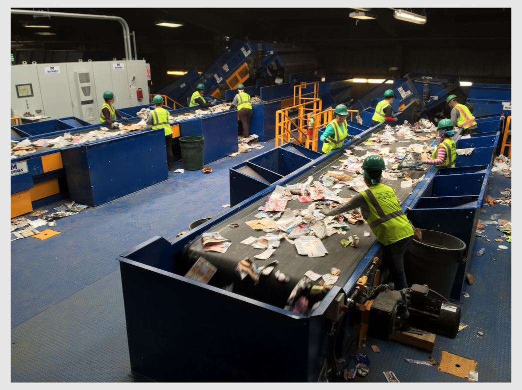
Lanqi Fei, Peixuan Huang, Zixiao Huang| MSiA 432 Deep Learning | Spring 2021 | Northwestern University

Problem Statement

The problem of our project addresses the recognition and classification of a trash object in a given image scene. Through this project, we intend to deliver a partial automation solution which facilitates the human efficiency of garbage collection and processing. Furthermore, this recognition functionality can be implemented in a fully automated robot for garbage collection.

In comparison to a traditional image recognition problem, the identification of a trash object is harder to solve because the categories are solely based on what the object is made of. So objects with similar shapes and colors can belong to totally different categories. It is rather challenging for the machine to reach a classification accuracy close to humans'.

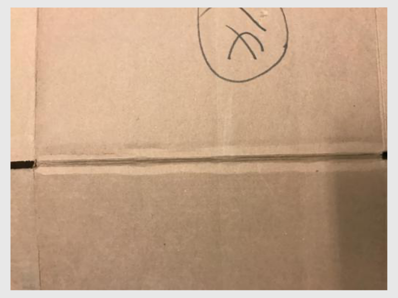
Our proposal approach to this problem is using a CNN model along with the performance tuning techniques learned from the MSiA 432 course.



Dataset

- Kaggle Garbage Classification Dataset and TACO Trash Dataset
  - ~ 2500 images total
  - 5 different trash classes (cardboard, glass, metal, paper, pastic)
- Data Cleaning and Challenges
  - The dataset initially has a class called "trash", which seems to be a combination of the objects of the other classes. Thus, we decided to eliminate the class and reassign the pictures in that class to other groups
  - The size of the dataset is relatively small
  - Some photos in the TACO dataset have more than one pieces of trash, thus we removed those from the dataset
- Data Augmentation and Modeling
  - We tried different methods of augmentation including randomly flipping the images, randomly rotating the images and shifting them both horizontally and vertically
  - The dataset is not very large so it doesn't require much computational power
  - Overfitting is not likely for the dataset
- Data Examples

Cardboard



Pastic

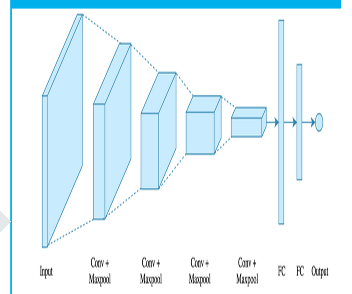
Glass

Technical Approach

Step 1

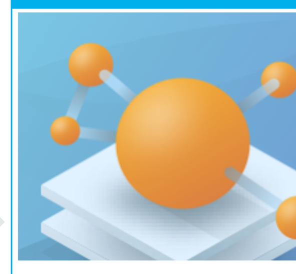
Data Preprocess

Step 2



Train CNN models and Analyze Outputs

Step 3



Parameter Tuning and Evaluate Results

Two CNN architectures are explored. Below are the parameter searching results.

VGG Network (56%)

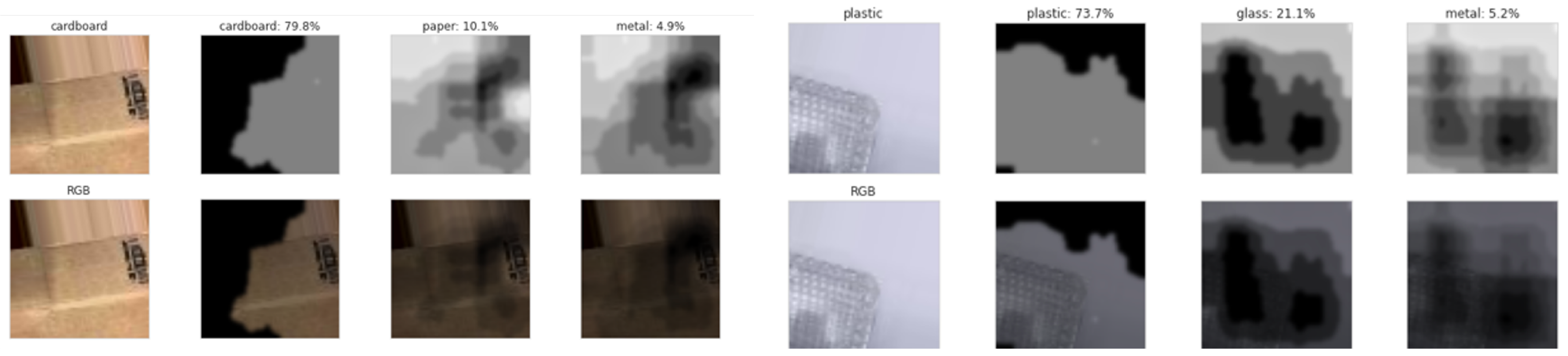
- Image Size: (64 x 64)
- Number of Blocks: 4
- Transfer Learning: 2 layers
- Augmentation: rotation, shifting, flipping are not particularly helpful
- Optimizer: Adam

Xception (82%):

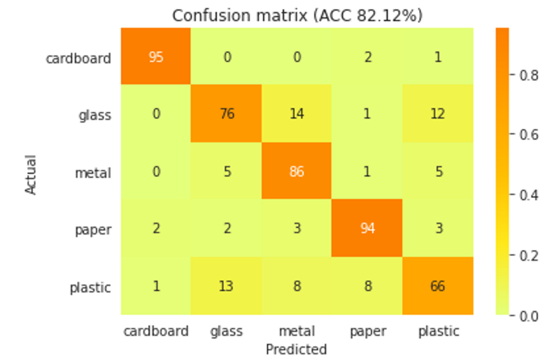
- Image Size: (64 x 64)
- Dropout Rate: 0.2
- Augmentation: random rotation, width and height shifts
- FClayer Size: 128

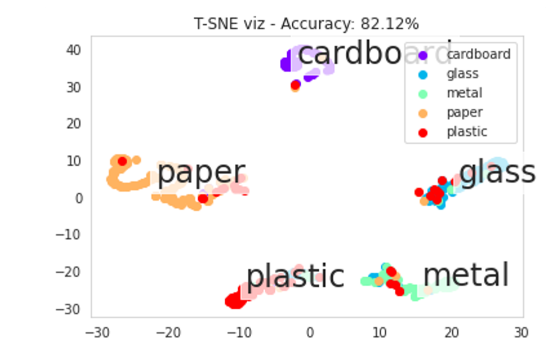
Heatmaps:

- Models search for patterns (e.g. brands, colors) to identify object class.

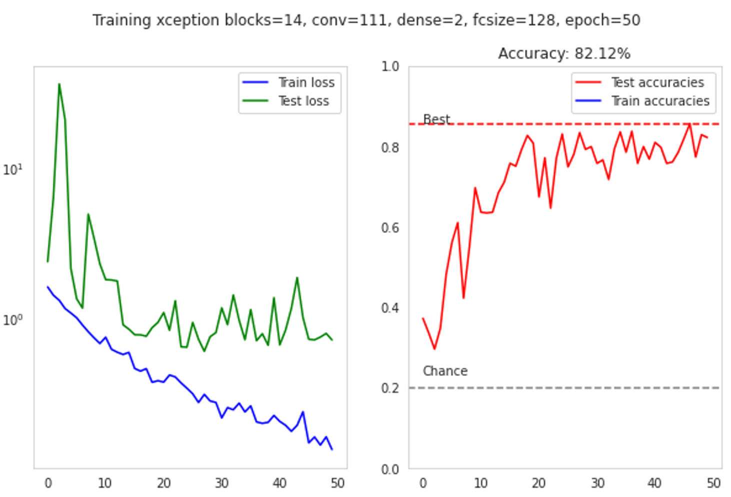


Results





- After running multiple models with different configurations, we obtained our best Xception model with an accuracy of ~85% at the highest and fluctuating at ~80%.
- The model did a plausible job on categorizing cardboards, metals, and paper. It struggled the most with plastics. There are several cases where the model misclassified the glass into metals or plastics and misclassified plastics into glass, metals, or paper.



- The mistakes our model made are fairly understandable. Glasses and plastics can be very similar and even hard for humans to identify when they are transparent. Some dark glass object can look like metals. Plastics and paper can also be similar. Overall based on the T-SNE graph, we saw a good separation, and the heatmaps suggest our model picked the right parts to identify in many cases.
- We estimate that a good human accuracy for this problem could be ~90%, so our model is not quite there. However,based on our observations and interpretations, we believe the mistakes our model made are highly human-like, and the performance is acceptable for a class project.

Conclusion

- Our model reached an accuracy around 80% and did an overall good job of classifying 5 major recyclable trash categories. The areas where the model is struggling are very similar to where a human would struggle, which is a good sign.
- Our approach is limited because we only applied 5 categories, and this may not be enough in certain industrial scenarios. Additionally, since our data images are mostly with neural backgrounds, the functionality would not adapt well to more diverse environments.
- The major improvement we can make in the future is to enlarge our dataset with more categories and more diverse backgrounds. This would help our model adjust to more general use cases. Another improvement would be to try other novel deep learning models which are not covered in MSiA 432 course.

References and Related Work

- Ceunen, Bouwe, "Garbage Detection with TensorFlow", https://www.kaggle.com/bouweceunen/garbage-detection-with-tensorflow
- Ceunen, Bouwe, "Training SSD MobileNet v2 with TACO dataset", https://www.kaggle.com/bouweceunen/training-ssd-mobilenet-v2-with-taco-dataset
- Dewage, Ranmal, "Transfer Learning Using VGG16", https://www.kaggle.com/ranmaldewage/transfer-learning-using-vgg16
- Gupta, Pranav, "Using CNN [Test Accuracy- 84%]", https://www.kaggle.com/pranavmicro7/using-cnn-test-accuracy-84
- Kuo, Richard, "Garbage CNN", https://www.kaggle.com/rkuo2000/garbage-cnn
- Singhal, Gaurav, "Importing Image Data into NumPy Arrays", https://www.pluralsight.com/guides/importing-image-data-into-numpy-arrays