CS230

Garbage Harvesting using Deep Learning

Goutham K T Muppala
Stanford University
krish24@stanford.edu

Abstract

Efficient Waste Management and recycling has been big challenge for many nations. our proposed machine learning model provides a way to automatically sort, there by reducing mixing of various categories of garbage at the dumpster which can lead to contamination of recyclable garbage. Transfer learning technique is used with ResNet18 and Resnet50 models which are pretrained on Image-Net[1] Data set. This pre-trained model is trained again with our data set to make predictions into various categories of trash. Instead of just using plain object classification we will also capture various features from the objects like CA CRV redemption, these features can help to improve accuracy further and also help to gain revenue from CRV redemption. our model achieved high test accuracy of 92.36% with resnet50.

1 Introduction

Garbage Recycling and Sorting manually has been a huge challenge. Manual mistakes of mixing various garbage categories leads to contamination of recyclable garbage at the recycling facility. As per CA recycle only 60-70 of recyclable garbage gets recycled rest goes to landfills. Main motivation for this project is to solve the problem in my apartment community where people fail to discard trash properly in the corresponding dumpster based on category. Is there way people could be rewarded for properly disposing trash by generating some value out of it? There seems to be smart trash bins by Indian startup[2] which rewards users for recycling bottles with 15 minutes of free WIFI. However here in USA we may find difficult to attract people with free WiFi,so we can use revenue generated by recycling trash or CA redemption's to give away free power ball tickets.

The purpose of the project is to create a deep learning model that can classify various types of garbage and also extract features from the labels or signs on the garbage like CA recycle, return credit that can potentially generate some revenue by directing these items to Recycling center. Our model takes image as input and provides classification material type of cardboard, glass, metal, paper, plastic and trash. This is further labeled by extracting any text features like UPC, recycle, reuse and CA CRV redemption. we will be using multiple models using the same data set to extract different features to achieve higher accuracy.

2 Related work

Previous related work for trash classification was done by prior Stanford students G. Thung and M. Yang for cs229 project [3] and Maoguo Shi, Qinyue (Jessica) Gu, Grace He for cs230 project [4] where they mostly explored object classification using CNN and SVM and they achieved best accuracy using SVM which is around 63% and 79% respectively on the same Trash Net data set [8]. Another work by Aral et al [5] on trash net data set using tranfer learning where they used variants of dense nets and inception nets to achieve 95% of validation accuracy. Another work by Yang et al[6] they proposed unified learning framework in which they utilized transfer learning and online learning. they created GarbageNet-101 which achieves 96% accuracy on their custom data set. However there is computation cost associated with the garbageNet-101 as it uses highest number of parameters for computation.

For scene text detection in natural scenes which we will use for identifying text on the labels of recycle products. Zhou et al [10] proposed EAST: An Efficient and Accurate Scene Text Detector where a fully convolution network adapted for text detection that outputs dense pixel predictions of lines or words followed by thresholding and non max suppression on predicted pixels. Baek et al[9] proposed CRAFT: Character Region Awareness for Text Detection where a fully convolution network based on VGG-16, they generate character boxes of text and also affinity boxes of individual characters. These character and affinity box scores are used to draw boundary for each word in the scene. we have utilized craft for label text detection.

3 Dataset and Features

we have utilized Trash Net data set for our object classification task. it consists of six different categories of items cardboard, glass, metal, paper, plastic and trash. I have gathered few more samples that will used for CA CRV text detection as existing data set have very limited images. Below is sample for the Trash Net data set.

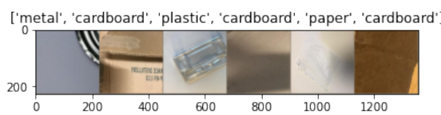

Figure 1: Various Trash categories

From Trash Net we will be dividing data set of 2527 images into training set of 1767 images and test set of 430 images and validation set of 327 images. Also we cropped images into 224x224 dimension during prepossessing. we also used horizontal flip, central crop and normalized input data set during prepossessing of data set.

4 Methods

For First milestone1, I have used resnet18 and resnet50 to train multi class object classification model. Resent's are pretrained on Image Net data set. we have used transfer learning technique to use the pretrained weights to initialize at start of training instead of randomly initializing the weights. we have trained our Trash Net data set on this pretrained weights with output layer of fully connected Linear output with six classes. we have used cross entropy loss function during our training.

$$\ell(x,y) = L = \{l_1,\dots,l_N\}^\top, \quad l_n = -w_{y_n} \log \frac{\exp(x_{n,y_n})}{\sum_{c=1}^{C} \exp(x_{n,c})} \cdot 1\{y_n \neq \text{ignore\_index}\}$$

Figure 2: Cross Entropy Loss

For Milestone2 we have trained same resnet18 and resnet50 with multi margin loss function and plotted various performance metrics for multi class object classification. Below is loss function for multi margin loss.

$$\text{loss}(x,y) = \frac{\sum_i \max(0, \text{margin} - x[y] + x[i])^p}{\text{x.size}(0)}$$

where $x \in \{0, \cdots, \text{x.size}(0)-1\}$ and $i \neq y$.

Figure 3: Multi Margin Loss

Above classification only classifies objects based on external features but it doesn't take into account various textual information present on them to get further insights into it. These textual information

Model	Loss	Best Test Accuracy	Best Val Accuracy	Test F1-Score
resnet18	Multi Margin Loss	89.78%	88.38%	0.89
resnet18	CrossEntropyLoss	88.84%	90.52%	0.91
resnet50	Multi Margin Loss	91.63 %	92.36%	0.92
resnet50	CrossEntropyLoss	91.40%	91.74%	0.91

Table 1: Evaluation Metrics for Garbage classification

will help to further classify glass and plastic so that they can recycled or returned at recycling facility for money.

In order to identify the textual features on the labels of the plastic and glass bottles we have to do a two step process where first we need to identify the text regions and the crop individual words or sentences to convert them into strings. For identifying the text regions we have used two different algorithms CRAFT[9] and EAST[10]. They both are used for text detection in natural scenes. we will using these techniques to identify text on the labels. Further Region of Interest (ROI) are cropped from the actual image. These cropped images are passed to tesseract OCR[11] tool with custom configuration to convert them to strings that helps to find if bottle/plastic container is CRV eligible. we had to employ cropped image rotation to improve the accuracy for vertical texts on labels.

5 Experiments/Results/Discussion

All our experiments are performed on Microsoft windows 10 Insider Edition with NVIDIA GeForce RTX 2060 GPU for training, test and validation. Pytorch is used for developing deep learning models.

We trained our model for 25 epoch's and used Stochastic Gradient Descent (SGD) optimizer with learning rate as 0.001 and momentum of 0.9 and also we have used learning rate decay at 0.1 for every 7 epochs. For resnet18 with Cross Entropy Loss we achieved validation accuracy of 90.82% and best training accuracy of 94.74 % . Similarly for resnet50 also we used the same hyper parameters and we achieved validation accuracy of 91.13 % and best training accuracy of 95.02%.

For milestone2 we have used Multi Margin Loss function which helped to improve the resnet50 accuracy by 1% at 92.36%. Table1 provides details best test and validation and F1-score for the classification of garbage. we have achieved highest F1-Score for resnet50 with Multi Margin Loss function. given the computational complexity with resnet50 which has approx 22 million parameter to compute, resnet18 with Cross Entroy loss provided better F1-Score of 0.91 with only 11 million parameters to compute i.e., half the computation cost.

For plastic or glass category label text detection we have used EAST[10] and CRAFT[9] text detection techniques, due to training complexity i have used the pretrained models provided by paper authors. For text area detection CRAFT performed better as it based on word affinity calculation and it was able to exactly identify the bounding for boxes for visible texts. Comparatively EAST missed single characters between the words. In our case CRAFT performed better for our use case. ROI identified by above image are passed to pytesseract to process and convert to string. we employed cropped image rotation to orient text horizontally which is when pytesseract performs text conversion better with higher accuracy.

pytesseract is also used to directly interpret text details but in most cases it only identified text that is straight and in lines which will not be the case for labels as they can be distorted and oriented in different directions. so we used pytesseract after identifying text ROI.

6 Conclusion/Future Work

we have provided models to classify the garbage with high accuracy 92% using transfer learning and also used combination of CRAFT and pytesseract to identify text on label of plastic or glass bottle to find out if they are CRV eligible. From different models resnet50 with Multi Margin Loss function achieved highest test accuracy of 92.36% and F1-score of 0.92. For Future work we can optimize the model to run using mobile nets this can be computed on edge devices. we can also implement incremental learning model where it can identify the object based on the textual feature on it instead of psychical features, there by system can learn by itself and over the time it can become more accurate.

References

[1] J. Deng, W. Dong, R. Socher, L. Li, Kai Li and Li Fei-Fei, "ImageNet: A large-scale hierarchical image database," 2009 IEEE Conference on Computer Vision and Pattern Recognition, 2009, pp. 248-255, doi: 10.1109/CVPR.2009.5206848.

[2] "Indian duo develops garbage bin that becomes Wi-Fi hotspot after it receives trash" https://www.planetcustodian.com/indian-duo-develops-garbage-bin-that-becomes-wi-fi-hotspot-after-it-receives-trash/8784/

[3] G. Thung and M. Yang. "Classification of Trash for Recyclability Status". CS229 Course Report, 2016, http://cs229.stanford.edu/proj2016/report/ThungYang-ClassificationOfTrashForRecyclabilityStatusreport.pdf.

[4] Maoguo Shi, Qinyue (Jessica) Gu, Grace He. "Trash Classification Using Convolutional Neural Networks". CS230 Course Report, Spring 2020 http://cs230.stanford.edu/projects_spring_2020/reports/38847029.pdf

[5] R. A. Aral, Ş. R. Keskin, M. Kaya and M. Hacıömeroğlu, "Classification of TrashNet Dataset Based on Deep Learning Models," 2018 IEEE International Conference on Big Data (Big Data), 2018, pp. 2058-2062, doi: 10.1109/BigData.2018.8622212.

[6] J. Yang, Z. Zeng, K. Wang, H. Zou and L. Xie, "GarbageNet: A Unified Learning Framework for Robust Garbage Classification," in IEEE Transactions on Artificial Intelligence, vol. 2, no. 4, pp. 372-380, Aug. 2021, doi: 10.1109/TAI.2021.3081055.

[7] Crawshaw, Michael. "Multi-task learning with deep neural networks: A survey." arXiv preprint arXiv:2009.09796 (2020).

[8] Garbage classification dataset https://www.kaggle.com/asdasdasasdas/garbage-classification

[9] Baek, Youngmin, et al. "Character region awareness for text detection." Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2019.

[10] Zhou, Xinyu, et al. "East: an efficient and accurate scene text detector." Proceedings of the IEEE conference on Computer Vision and Pattern Recognition. 2017.

[11] Tesseract OCR https://github.com/tesseract-ocr/tesseract