Mitacs Project Proposal: Deep Learning Dataset Augmentation for Detecting Rare and Severe Contaminants in Waste Management Collection

Nolan Flegel, Rishabh Prasad, William Peers

Submitted: November 11, 2021

Revised: January 16, 2022

Table of Contents

Abstract

Mismanagement of recyclable waste is an environmental disaster and a burden on society. Recycling is a costly, risky and inefficient process. This project's purpose is to increase recycling collection rates and reduce operator costs by identifying rare and severe contaminants at the source. This project has two goals: (1) to create a universal waste bin classification model and (2) to generate datasets for rare and severe contaminants using deep learning techniques.

The first task is to develop a Universal Waste Bin Detector using Image Segmentation and Fully Convolutional Neural Network (FCN). This task will utilize transfer learning with the existing binary classifier to develop a universal segmentation model for waste bin detection. The universal model will be evaluated using existing datasets for true positive, false positive and true negative samples.

The second task is developing a Copy-Paste Augmentation pipeline. Annotated image masks of known rare and severe contaminants will be inserted into our industry partner's existing datasets using the Simple Copy-Paste Augmentation Algorithm to generate new datasets. Using Mask R-CNN, detection models will be trained using the augmented dataset and tested against an annotated validation dataset. The goal of this task is to test the validity of artificially generating unique datasets for rare events and evaluating the model performance against real world occurrences.

Background

Our industry partner is an Artificial Intelligence company offering solutions for waste management object classification. Our industry partner's existing technology captures all data on the remote recycling collection vehicle before uploading to the cloud for analysis. Machine Learning is used to identify and classify contaminants found in waste bins. Our industry partner has partnered with multiple Western Canadian municipalities and waste management companies to reduce recycling contaminants in residential collection programs. Our industry partner's Recycling Education Platform helps municipalities, waste producers and waste management organizations understand recycling contamination and waste generation at the point of collection.

With recent reports indicating that an abysmal 9% of plastic waste in Canada is collected and recycled, the goal of this project is to enable waste management firms to drastically improve plastic waste collection by removing contaminants and reducing overall solid waste from entering landfills (Deloitte Canada, 2019). Expanding upon our industry partner's existing contaminant datasets by using copy-paste augmentation to create new masks to identify harmful materials such as yard waste, propane tanks, needles, diapers and batteries in recycling bins. These items are rarely found in publicly available datasets such as TrashNet and TACO which contain images of contaminants found in recycling; however, it is critical to identify these rare occurrences (Proença & Simões, 2020; Aral et al., 2018). TrashNet [https://github.com/garythung/trashnet] is a collection of annotated images with a single object present on a white background with even lighting (Thung & Yang, 2017). It contains approximately 2500 images of recyclables with only a small amount of trash (contaminants). This dataset is helpful for single object classification but it would not be applicable for training multi-object classification models. Since it can't be extended for multi-object classification, it can be used as basic test data for measuring the generalization of a model. In addition, the creators of the TrashNet dataset also developed two classifiers: an eleven layer Convolutional Neural Network (CNN) that is very similar to AlexNet, and a support vector machine (SVM) with a gaussian kernel using the SIFT algorithm for feature extraction. The SVM achieved better results than the CNN. It achieved a test accuracy of 63% using a 70/30 train/test data split. While the training error was 30%, the SVM is a relatively simpler algorithm than the CNN (Thung & Yang, 2017). Although research in this area can give direction for model development, they have to be improved to classify objects to be useful for classifying curbside waste (Adedeji & Wang, 2019).

In this project, data collection is achieved through outfitting the collection vehicles with cameras and computers to detect waste bins and collect data before uploading to cloud services. A binary detector is used to trigger image captures when waste bins are detected in the camera feed. Each vehicle platform is unique in that the hopper dimensions, hopper orientation, camera placement and camera field of view will vary between vehicles. Currently at our industry partner, bin classification training is unique for each vehicle and training must be completed before that platform can begin operation.

The significance of a universal model for detecting recycling bins is a crucial component to rapidly scale the number of municipalities the technology at our industry partner impacts. The existing binary waste bin classifier has a number of limitations; each vehicle requires a uniquely trained model and is unable to classify a variety of waste bins types. These bins are often different shapes and sizes contingent on the company or municipality they are associated with; however, they share similarities which create an opportunity to improve the installation process tremendously. This universal model of detection would allow bin detection to be seamless regardless of the municipality or company the waste bin belongs to.

Copy-Paste Augmentation is proving to be an effective method for building accurate instance segmentation models using fewer images (Ghiasi, et al., 2021). Using COCO datasets, the authors of Simple Copy-Paste showed significant improvements in model performance. This paper also showed that the copy-paste method can be used for supervised and unsupervised training and that self-training annotations may be more accurate than human annotation (Zoph, et al., 2018). This project will use copy-paste techniques to develop synthetic datasets for rare and severe contaminants. These rare occurrences are essential to detect as they pose direct risk to the operator and downstream personnel. This research will show that it is possible to generate accurate segmentation models for training when real world data cannot be collected.

Objective

The objective of this project is to develop a universal machine learning model for detecting curbside collection events and the creation of a pipeline for generating augmented datasets of rare and severe materials. The first task is to create a waste bin detection and classification model that is agnostic of the vehicle and hardware platforms. A dataset of images collected from our industry partner platform will be separated into different categories, image masks for each category created, and instance segmentation models trained.

The second task is the development of an automated copy-paste augmentation pipeline which merges images containing annotated rare contaminants with our industry partner's existing waste dataset collected from recycling vehicles. The pipeline performs various synthetic transformations on the image to create a large dataset with different lighting, rotation, cropping, and other image manipulations to form varied and diverse datasets. These new datasets can be used to train models that can detect and classify rare and severe contamination in waste collection.

Specific Objectives

(1) Universal Waste Bin Detector:

- Consolidate original binary classifier training data on waste bins from various municipalities to be used as part of the instance segmentation training datasets. Review images captured by the existing binary classifier and identify false positive images for segmentation categorization.

- Categorize and create annotated image datasets of positive and false positive images of waste bins.
- Use transfer learning and image segmentation to train a new detection model.
- Test and benchmark new universal waste bin classifiers using false positives and true negatives from our industry partner's existing datasets.
- Deploy the universal bin classification model and evaluate its performance in the real world.

(2) Copy-Paste Augmentation Pipeline

- Collect a small number of sample images and create image annotation masks for rare contaminants that will be used as the foundation of the new synthetic dataset.
- Develop an augmentation pipeline where the Copy-Paste algorithm is applied to the annotated image masks and merged with our industry partner's existing datasets to create a new synthetic dataset.
- Use the new dataset to train a Mask R-CNN instance segmentation model.
- Prototype dataset pipeline using a small subset of the dataset.
- Analyze the dataset generated from copy-paste augmentation and generate a balanced test dataset to establish minimal opportunities of introducing bias into the model during training.
- Scale prototype pipeline to utilize the entire dataset.
- Train Mask R-CNN model on the test dataset containing rare contaminants.
- Evaluate the detection model using annotated true images from existing waste datasets.
- Deploy rare contaminant model to industry partner's platform and evaluate performance under real-world conditions
- Evaluate model inferences on existing dataset to identify previously undetected items of interest
- Iteratively incorporate detected items into the instance segmentation model.

Methodology

Universal Waste Bin Detector

To achieve a generic model which can be universally applied, our industry partner's datasets of waste bins from various municipalities will be consolidated. Upon this existing data, transfer learning will be applied to develop a universal model. Transfer learning takes advantage of existing models which have been trained with a specific dataset by transferring the knowledge the artificial intelligence has gained during its training on an additional but similar dataset (Weiss, et al., 2016). Instance segmentation will be used to classify different types of waste bins as well as the distance the bin is located from the waste collection vehicle. There are several competing platforms and algorithms for real time image segmentation. Single pass methods such as YOLO have been shown to have the best performance on real-time instance segmentation (Buric et al., 2018). Two algorithms will be compared and evaluated for performance on the edge computer and a USB Tensor Processing Unit (TPU).

- YOLOv5 (Buric et al., 2018)
- YOLACT++ (Bolya et al., 2020)
- Intel Neural Compute Stick 2

Both YOLOv5 and YOLACT++ are extensions of the YOLO object detection method and represent an iterative improvement upon the method (Bolya et al., 2020). These algorithms were selected for their real-time object detection capabilities and the following process will be used to develop and evaluate their performance:

- Historical images from the collection vehicles will be collected, processed and categorized for false positive bin detections.
- Create image dataset by combining the false positive images with existing true positive and true negative datasets.
- Create instance segmentation models using the categorized dataset.
- Development of an evaluation module where the algorithm performance can be measured against the previous binary detection model
- Comparison of frame rates and processing delay between the new universal model and the existing binary detection model
- The model will be deployed to the edge computers installed on waste collection trucks where performance can be tested with real world events.
- Processing will be offloaded to a Tensor Processing Unit (TPU) plugged into the edge computer.
- Compare FPS performance between the TPU and the edge computer CPU.

Copy-Paste Augmentation Pipeline

Detection of rare or severe contamination is difficult, however important to identify as these events can cause great harm to employees, the general public or infrastructure. Datasets containing these rare objects are difficult to obtain, therefore this project will explore methods of generating and augmenting synthetic datasets. This progressive approach reduces the overhead required to collect and process raw images, resulting in lower processing time and costs associated with training detection models. For example, it required 22 worker hours in order to create 1000 instance masks in the COCO dataset, this time can be saved using the Copy-Paste technique (Ghiasi, et al., 2021).

Copy-Paste Augmentation enables the creation of large datasets using a few images via synthetic transformations. This process can be automated, accepting a few example images of a new contaminant and producing a dataset that can be used to train the model using Mask R-CNN. Models trained with synthetic data generated from the Copy-Paste algorithm have proven to be a robust solution which perform similarly or better when compared to traditionally trained models (Ghiasi, et al., 2021). This augmentation method is also actively evaluated against other state-of-the-art augmentation methods using COCO and LVIS benchmarking standards to enhance its performance.

The industry partner requires an object classification solution that can identify many partially obscured objects. Mask R-CNN will be used to provide instance segmentation masks for each object. When generating new synthetic datasets, the augmented objects will need to be integrated into the existing datasets which have objects layered and partially obscured by material within each image. This is not possible with bounding boxes and instance segmentation is required. While other models such as YOLACT++ may yield faster detection times, the Mask R-CNN model was shown to have more accurate instance segmentation results on smaller objects. The industry partner's application prioritizes accuracy over speed, therefore Mask R-CNN is preferred (Bolya et al., 2020).

The method Cascade Eff-B7 NAS-FP is known as Simple Copy-Paste evolves a number of augmentation techniques which significantly improved performance when compared to other implementations of cut-and-paste augmentation:

- Using large scale jittering (LSJ) with an image resize range of 0.1 to 2.0.
- Self-Training using ground truth instances.
- Random horizontal flipping.
- Not implementing geometric transformations.

The proposed augmentation pipeline will implement the Simple Copy-Paste method using our industry partner's' waste collection dataset and annotated sample images of a new object for detection. Randomly selected images from the sample dataset will be pasted into existing images from our industry partner's waste collection dataset. Random image transformations such as jittering and horizontal flipping will be applied to each image, followed by taking subsets of the images and inserting copies into the other images. (Ghiasi, et al., 2021). Our implementation of the Simple Copy-Paste method will require new image masks to be layered and obscured by existing masks. The objective is to generate a dataset that encapsulates the real world representations of the desired objects. The Simple Copy-Paste method provides useful transforms and augmentations can be applied for this purpose. The transformations applied to each image including the synthetically pasted objects will require extensive applied research to generate a new reliable dataset to train a machine learning model. Although random transformations are applied, they must be adjusted, modified, tested, and validated exhaustively to ensure the model is trained to identify new objects effectively. The resulting instance segmentation model will be validated using real world data collected from our industry partner's platform. After benchmarking model performance on the validation data, the model will be applied to the industry partner's historical collection events to detect and classify previously undetectable objects. These new objects can be incorporated into the model, further enhancing the performance.

The Copy-Paste pipeline implementation will follow this process:

- Collect Sample Images of the new object and create instance segmentation annotations.
- Split real-world dataset appropriately for training, testing, and validation.
- Import our industry partner's existing dataset of waste collection images.
- Import sample images of the rare contaminant.
- Utilize segmentation masks of the new object to crop, apply further transformations, and paste objects into the existing dataset.
- Generate a dataset of images with layered annotations of both the existing waste collection dataset and the new object. Layered objects must overlap and be partially obscured by other masks.
- Train a Mask R-CNN model using the synthetic dataset.
- Benchmark performance using real-world data collected from our industry partner's platform.
- Adjust and modify transformation parameters to improve the dataset's effectiveness if required
- Deploy the model to our industry partner's cloud services where real-world instances of waste collection will be processed
- Compare performance of detecting the new object as well as other previously defined classes.

References

Aral, R., Keskin, Ş., Kaya, M., Hacıömeroğlu, M. (2018). Classification of TrashNet Dataset Based on Deep Learning Models. 2018 IEEE International Conference on Big Data. https://doi.org/10.1109/BigData.2018.8622212

Bolya, D., Zhou, C., Xiao, F., Jae Lee, Y. (2020). YOLACT++: Better Real-time Instance Segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence 2020 Aug 5. https://doi.org/10.1109/tpami.2020.3014297.

Buric, M., Pobar, M., Ivasic-Kos, M. (2018). Ball Detection Using Yolo and Mask R-CNN. 2018 International Conference on Computational Science and Computational Intelligence (CSCI). https://doi.org/10.1109/CSCI46756.2018.00068.

Deloitte Canada. (2019) Economic Study of the Canadian Plastics Industry, Market and Waste. https://publications.gc.ca/site/eng/9.871296/publication.html.

Dvornik, N., Mairal, J., Schmid, C. (2018), Modeling Visual Context is Key to Augmenting Object Detection Datasets, 15th ECCV 2018: Munich, Germany - Volume 12. https://arxiv.org/abs/1807.07428.

Ghiasi, G., Cui, Y., Srinivas, A., Qian, R., Lin, T., Cubuk, E., Le, Q., Zoph, B. (2021). Simple Copy-Paste is a Strong Data Augmentation Method for Instance Segmentation. arXiv, https://arxiv.org/abs/2012.07177v2.

Proença, P., & Simões, P. (2020). TACO: Trash Annotations in Context for Litter Detection. arXiv. https://arxiv.org/abs/2003.06975

Jainuddin, A., Hou, Y., Baharuddin, M., Yussof. S. (2020). Performance Analysis of Deep Neural Networks for Object Classification with Edge TPU. 2020 8th International Conference on Information Technology and Multimedia (ICIMU). https://doi.org/10.1109/ICIMU49871.2020.9243367.

Weiss, K., Khoshgoftaar, T.M. & Wang, D. (2016). A survey of transfer learning. J Big Data 3, 9 (2016). https://doi.org/10.1186/s40537-016-0043-6

Zoph, B., Ghiasi, G., Lin, T., Cui, Y., Liu, H., Cubuk, E., E., Lej, Q. Rethinking Pre-training and Self-training. (2018) 34th Conference on Neural Information Processing Systems (NeurIPS 2020), Vancouver, Canada. https://arxiv.org/abs/2006.06882v2.

Adedeji, O., & Wang, Z. (2019). Intelligent Waste Classification System Using Deep Learning Convolutional Neural Network. Procedia Manufacturing, 35, 607-612. doi:10.1016/j.promfg.2019.05.086

Thung, G., & Yang, M. (2017). Garythung/trashnet. Retrieved October 28, 2020, from https://github.com/garythung/trashnet