Team Copy-Waste - Project Bazaar 1 Script

Team Member (Re)introductions

- Nolan - Machine Learning Lead
- Will - Back-End Services Lead
- Rishabh - Front-End Services Lead

Vision:

Our project has two objectives; reduce risks to the public, waste management employees and facilities and reduce the cost to municipalities and waste management companies caused by rare and severe contaminants in recycling waste collection.

Mission

We strive to achieve our vision by automating the detection of severe and rare contaminants in the recycling stream

Business Need

Our business need, to reiterate, is to reduce risk and cost to municipalities and waste management employees. As recycling rates increase, contamination and risk becomes a growing concern.

Fires are a major concern at waste facilities: Here in Regina there were two serious incidents in 2021. This is suspected to be caused by batteries, which is currently not detectable as there is no image dataset available.

Current: Status Green

Project issues / changes

- Limited resources on implementing copy paste augmentation leading us to rely heavily on personal research and experimentation as this methodology was published June of 2021.
- TPU stick for processing tensors is no longer necessary as our Industry Partner has upgraded their edge devices to a more powerful model.
- First prototype for Copy-Paste will use yard waste images as we have a readily available source of validation images, whereas validation data for propane tanks and batteries requires cooperation of a third party.

Addressing Comments

Craig’s Comments

- To address Dr. Gelowitz comments regarding IP, we are now in the process of signing a contract with our industry partner.

Adam's Comments

- Regarding Adam's comments, we have allocated a 4 week period of time for testing and deploying the ML model produced by our copy paste pipeline.

Individual Contributions (Dec 6 - Jan 25):
Will:

- Collecting and annotated True Negative image datasets
- Separated and categorized images for testing and training using python scripts
- Created, trained, and generated results for initial YOLO detection model
- Creating python scripts to import an audited annotation dataset for Copy Paste

Rishabh:

- Collecting annotated images of audited hopper images
- Research implementations of Copy Paste data augmentation
- Begin prototype version copy paste augmentation on a subset of images

Nolan:

- Trained Bin Detector in Yolov5
- Evaluated Model Performance
- Collected additional training images and retrained detection model
- Created simulation videos to evaluate model performance with business logic.
- Working with our Industry Partner to deploy our model in their production environment and edge computers (recycling trucks).

Bin Tip Detector

This is a demonstration and comparison of our Bin detector between our first model and our most recent model. On the left you can see the original model struggled with classification and would frequently get confused. On the right, the re-trained model using more cycles, additional images and fewer classes, performs significantly better.

Here is the confusion matrix produced after training our most recent model, it has high precision and recall. We compiled a number of tables to evaluate model performance on validation and training dataset where we can see the number of false positives and false negatives. Our most recent model has an accuracy of 94.8% for detecting blue recycling bins. We tested our detector with a video containing 30 collection events, at 24fps our model detected 30/30 and at degraded performance of 4fps, the detector detected 28/30 events.

This past week we worked with our industry partner to deploy our bin detector to one of their test vehicles. The recycling truck operated a regular collection route over the course of a week and our detection model was evaluated for real-world performance. The initial results of the real world test are promising, our bin detector inferences on the edge computer were 300ms, which was within performance parameters and collection events were accurately recorded. The model performed as expected and will be deployed to additional vehicles and performance will be monitored over the next few weeks.

Simple Copy-Paste

This flow chart describes our intended plan for the data augmentation pipeline. It begins by a Waste Management worker identifying a contaminant for detection, then using the copy-paste data augmentation methodology, we can produce an artificial dataset composed of real world images from our industry partner and samples of the requested contaminant. The final product will be a trained object classification model.

Our current data augmentation prototype utilizes a background image and an image where objects will be extracted from. In this case the left image will be the background image and the right image consists of contaminants which will be extracted. Our goal is to paste these objects on the background with random transformations such as resizing, flipping, rotating, and layering with objects within the background image. This variety will allow us to generate an augmented image dataset which can be used to train a machine learning model. **Our** prototype was able to output the following image as a result of applying the simple copy-paste methodology which performs basic image transformations and applies the annotations.

Our industry partner has a massive available dataset and to assure optimal model training we wanted to make sure we were working with a “clean” dataset. Our industry partner has audited their dataset, leaving us with a list of identification numbers associated with “clean” images and annotations. In order to only use this fraction of the dataset, a python script was created to pull the audited images as well as generate a new annotations file for said images.

Waste Management Dashboard

This is our dashboard for municipal decision makers with easily accessible information on how various regions around the city are performing. Through coloured signifiers it provides information regarding performance such as number of offenders, contaminants detected and contamination rate.

Knowledge Management Overview

We have continued to perform weekly agile sprints, while managing our kanban board and recording meeting minutes to ensure we are progressing effectively.

Additionally, Mitacs reviewed our project’s scope and requested changes to further strengthen our project objectives. Accordingly, we adjusted our timeline, created more specific tasks as requested by the reviewers, and have submitted a revised document.

Next up

Team

- Meetings planned with our Advisors and with our industry partner
- Monitoring the deployed bin detector on the truck's edge computer

Rishabh

- Creating scripts for pasted object placement and transformations

Nolan

- Will work on Deploying Bin Detector and Layering augmented annotations.

Will

- Working towards running copy-paste on the audited dataset
- Working on exporting pasted images from the Copy Paste algorithm with combined annotations files

Team Reflection

- **Does the team feel "on track"? (reiterate the above colour status)**
    - We believe we are still on track
- **What progress does the team particularly feel good (great) about?**
    - We feel really great about the progress of the Bin Tip detector, we made significant improvements to the model's performance and our industry partner has enough confidence in our detector to deploy it to their production environment.
    - Our initial progress on Copy Paste has been extremely encouraging and we are excited about the next stages of our project.
- **What barriers (if any) does the team feel is a current impediment to success?**
    - Limited documentation has required us to allocate time for experimentation.
- **What help (if any) does the team require to move positively forward?**
    - Moving forward we need to keep having regular meetings with our industry partner to guide us.
- **What questions or concerns does the team have (if any)?**
    - No concerns at this time