Team Copy-Waste

Nolan Flegel, Rishabh Prasad, William Peers

Team / Instructor Scrum - 7

Mar.8, 2022 - Mar.15, 2022

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

No issues / changes at this time

Individual Contributions (Mar 8 - Mar 15):

Will:

- Exporting transformed images and annotations files to AWS
- Began auditing False Positive images of grocery bags
- Refactoring the code base to have standardized variable/function names

Rishabh:

- Audit False Positive Images of Clamshell packaging
- Implement Importing / Exporting images using organization standard
- Scaling up dataset production and experimentation
- Implementing various error handling functionality so that images which have an issue during augmentation are rejected.

Nolan:

- Generated small scale datasets
- Testing and fixing issues that arose during large scale testing.
- Added additional error handling for multiple contours
- Experimented with data transforms options
- Produced large scale dataset (10,000 images)
- Began training Mask R CNN model using augmented data
- Audited false positive images for clear plastic packaging.

Simple Copy-Paste Overview

Our data augmentation pipeline has been scaled up. We performed several tests using real data and added features to solve any problems discovered. Our pipeline was able to process a large volume of data without any hard runtime failures. We were able to produce several artificial datasets using sample images of yard waste. We started by producing 100 image datasets where we tested various configurations and different transformations. We verified that our outputs were working correctly. Our current dataset consists of 10,000 augmented images stored on an aws s3 bucket and we are training detection models with it.

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. Our focus is now also shifting to complete project completion documents.

Next up

Team

- Code Reviews
- Model training and testing
- Project Conclusion documentation

Nolan

- Training Mask R CNN models using augmented datasets
- Generate performance metrics for resulting Models
- Continue integration with industry partner and set up automated retraining of models
- Code reviews and documentation.

Will

- Merging artificial annotations files with existing ones to expand and diversify our datasets
- Code reviews and documentation
- Assisting with work on the Dashboard

Rishabh

- Scripts to evaluate performance of augmented model
- Implementing dashboard changes as identified with user testing
- Assist with model training where necessary

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 great about producing large datasets. We were able to scale up our pipeline to handle a variety of incoming data sources. Our method is able to handle runtime errors gracefully, logging any errors and the corresponding image while continuing to process the remaining images.
- **What barriers (if any) does the team feel is a current impediment to success?**
  - Potential for unexpected problems with model training.
- **What help (if any) does the team require to move positively forward?**
  - Guidance and resources from our industry partner regarding Sagemaker and Mask R CNN training.
  - Do we need to book a demo space?
- **What questions or concerns does the team have (if any)?**
  - We are concerned about potential problems with model training and performance. We know there are a number of unknown factors related to experimental/ cutting edge software development and we have created procedures to mitigate certain risks. Our project is reaching a critical point where all of our efforts over the past 6 months are coming together and justified.