University of Regina

After Action Review

Copy-Waste

Nolan Flegel, Rishabh Prasad, William Peers

Published: April 8, 2022

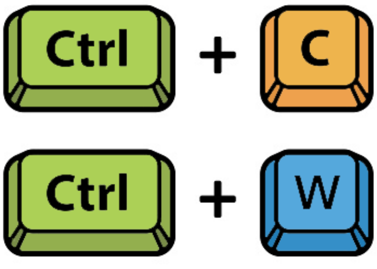

Table of Contents

What Was Expected to Happen?

Initial Purpose and Objectives?

The purpose of this project was to iterate upon Prairie Robotics smart recycling platform. The objective being to expand upon their current technological stack to decrease the environmental impact caused by residential recycling, as well as reduce risks and costs associated with it. To uphold this requirement, we collaborated with Prairie Robotics to identify a gap which existed in their current technology. Through conversations, we identified the problem of detecting rare contaminants such as propane tanks and batteries which pose serious risks. Thus, the primary goal became to detect rare objects which were previously undetectable. Additionally, we wanted to improve the scalability of their platform by implementing a universal bin detector which could be operational in any municipality. As a result, our collective goal was to be able to make municipal recycling simple, risk-free, and efficient.

We also intend to submit a proposal for our project to Mitacs to receive funding in order to conduct our objectives. In the previous year, Prairie Robotics alongside a team of students were successful in their submission, which ensured we could also receive similar compensation for our work.

Our personal purpose in conducting this project was to implement a software solution which has a positive impact on the environment. As young engineers, we find it critical to contribute to platforms which serve towards a greener future. It is important for us to think creatively and find innovative solutions using cutting-edge technology for the complex problems which exist in our environment. Achieving this objective required us to ensure we were working on a well defined, structured, real world. Collaborating with an industry partner provided us with a proper structure and experience for conducting sprints, code quality, code reviews, and business practices.

Our Audience

Our primary clients were Prairie Robotics Inc. and municipal waste management services, who would be using our product. Our capstone instructor Dr. Timothy Maciag and our capstone mentor Dr. Mohammed El-Darieby would oversee our progress and provide guidance and feedback where needed. The University of Regina and Mitacs were also stakeholders, as we entered into a partnership to receive academic funding for our project. Each stakeholder provided valuable insight and requirements that were incorporated into our project.

The members directly involved in the project would be the capstone group members, William, Rishabh and Nolan. Sam Dietrich, Matthew Barriault and Steven

Mikha from Prairie Robotics will provide technical guidance and other support for the project.

Initial Timeline

Our initial timeline was documented in a Gantt chart, as required by Mitacs. September would be spent doing research and completing the initialization phase of the capstone project, where we would learn about the expectations and requirements. Starting in October we would gather data and begin working on the universal bin detector. We expected to complete this first task around the first week of November and have a working object detection model. The experience and lessons learned from this initial product would lead into our work on the Artificial Data Augmentation Pipeline.

November and December would be spent exploring and experimenting with the Copy-Paste method as well as building out the architecture for the augmentation pipeline. By mid January we planned to produce our first artificial dataset and immediately start training a machine learning model. February and March were allocated for iterative improvements to the datasets and models, adding more artificial objects and testing the model performance. The project should be complete by the first week of March, allowing ample time to complete our project report and final documentation.

Facilitators and Barriers

We identified a number of barriers to success, the most significant being a lack of knowledge and resources for the Copy-Paste method. This is an emerging, cutting edge field and there are very few publicly available knowledge sources for this technique. Additionally, this is an unproven method and we cannot be certain of its success. Initial results from the Simple Copy-Paste research paper was promising and is the basis for our project, however there is uncertainty over our ability to implement these techniques in a meaningful way.

Our facilitators of success will have access to Prairie Robotics staff and platform. We will be iterating upon the work of the previous capstone group, Ray, Avery and Noah and the rest of the platform. The code base will provide a great template for how to structure our code and integrate our services. We will be able to leverage existing technologies and architecture, such as AWS and SuperAnnotate. The company has grown in size in the past year and their knowledge base has expanded. Prairie Robotics will be part of our weekly sprint planning sessions and they will be able to guide us through the project by helping us to prioritize tasks and simplify problems.

Intended Deliverables

Our envisioned products would be a waste bin object detection model and an artificial data augmentation pipeline. The object detection model would identify a variety of municipal waste bin styles and common false positives, such as Porta-Potties or Water Barrels. The augmentation pipeline would be a more complex system. It would take a few basic images of an object and using techniques outlined in a research paper, output a large, varied dataset for that object. The objective was to take rare objects found in residential recycling and create new datasets that could be used to train an object segmentation model. This model could be used on incoming recycling collection to detect rare and severe contamination. The long term goal is to detect these objects at the source and prevent them from entering the recycling stream.

What Actually Happened?

Actual Purpose and Objectives

We were able to develop solutions as we had initially planned. Although the objectives were ambitious, we were able to achieve all of our intended goals through proper planning, structure, and communication between the team. We developed a data augmentation pipeline which is able to detect contaminants which were undetectable previously. We also implemented a universal bin detector which Prairie Robotics was able to immediately apply to scale their platform. Lastly, we also developed an interactive dashboard for decision makers to leverage familiar knowledge and learn more about their recycling program. In addition, Mitacs approved our project objectives, enabling us to receive funding to accomplish our goals and cover the costs of this project.

Actual Audience

Our audience mostly remained the same as initially planned; however, we received the opportunity to bring Trevor Douglas as an additional academic mentor in the Winter 2022 semester. Prairie Robotics and municipality workers remained our primary stakeholders through our development process. We were able to use Prairie Robotics as an intermediate source to understand our end users problems and the experiences we wanted to develop. Our industry partner was also able to provide us with extensive guidance on challenges we faced with machine learning and code quality in general. On the academic side, we relied on Dr. El-Darieby, Dr. Maciag, and Trevor Douglas for their expertise in their fields of interest. We were able to rely on Dr. El-Darieby for support regarding Mitacs funding and his Ph.D student George Daoud for machine learning questions. We also received advice from Dr. Maciag for our user interface to ensure it follows people-centered design practices, but also regarding making our project understandable to a less technical audience. In addition to this, we reached out to Trevor Douglas who ensured our project was progressing according to the faculty's requirements.

Actual timeline

The timeline we outlined in our Mitacs proposal was reviewed and some revisions were requested. They suggested that we allocate more time for our bin detector, which turned out to be an accurate estimate as the bin detector took longer than we had initially planned. We also allotted more time to the data augmentation pipeline, specifically for testing and evaluating our output. Lastly, we included the development of a front end knowledge dashboard that could bring the whole project together.

Overall, we stayed on track for most of the project, falling a few weeks behind schedule at different points during the semester, only to catch back up soon after. Our weekly sprint planning sessions helped us prioritize effectively and stay relatively close to our schedule. That being said, there were a number of unexpected challenges and obstacles that caused problems with our initial timeline.

The mitacs project proposal was far more involved and complex than we expected. The research requirements and initial documentation required a significant amount of time and consultation. This pushed our starting dates back to October, at which point we were able to finally begin working on the universal bin detector. To further complicate our progress, we found that the provided dataset for the waste bins was insufficient and required additional time to collect and annotate new images. Once we had a high quality dataset to work with, progress on the bin detector moved rapidly.

By the end of November we had caught up to our initial schedule and things were looking promising. At this point, each of us were working on separate tasks in parallel; testing the bin detector, completing the dashboard and starting the data augmentation architecture. Something that caught us by surprise was the success of the bin detector. The performance statistics were so promising that Prairie Robotics was interested in deploying it as soon as possible. Completing the testing and transferring the project to Prairie Robotics added a few extra weeks of work and delayed some of the progress on the data augmentation pipeline.

In January, we again found ourselves a few weeks behind schedule, however we completed the bin detector and handed off management of the project to Prairie Robotics. We turned our full attention to developing the data augmentation pipeline. Initial progress was going well, however we quickly ran into some challenges when trying to use an existing implementation of the Copy-Paste method. There were a number of incompatible or incomplete features. We made the decision to start from scratch and implement our own version. This was a major, unexpected setback and resulted in significantly more work than we had planned.

With all of our attention on implementing Copy-Paste, we slowly started catching up to our envisioned schedule. By mid-February we were looking to be back on track and ready to produce a dataset over reading week. Unbeknownst to us, a serious data quality issue had been discovered in the Prairie Robotics dataset. There were a number of significant annotation inaccuracies and the vast dataset needed to be audited before we could begin augmentation. In conjunction with everyone at Prairie Robotics, we spent our next two weeks auditing a section of the dataset, which amounted to 30,000 images each.

Moving into March, we are nearing the end of the project. We are a little behind schedule because of the aforementioned auditing and this causes us to scale back some of our final deliverables. We are able to start testing our data augmentation pipeline and work on debugging any hard faults and unexpected behaviors. We produce our first couple artificial datasets and look to scale up to a sufficiently large set that can be used for machine learning. The only significant change to our deliverables is that we won't be able to gather ground truth data for propane tanks or batteries. This would require coordination with the City of Regina and would need to coincide with scheduled maintenance for one of their recycling trucks.

We are able to use some existing ground truth data for yard waste to complete the project within the timetable. The last few weeks of March are spent training a machine learning model to detect leaves. We tested the model against some real, unaltered images of recycling and were able to detect a number of new instances of leaves that were previously undetected. The last month was extremely challenging as everything started to come together. It was difficult to balance the workload of trying to complete the technical components as well as finalizing all the documentation and the presentation.

Products Produced

We produced three separate products: the Green Screen Dashboard, a deployed universal waste bin detector, and a data augmentation pipeline for detecting rare and severe contaminants.

The universal bin detector iterates upon a previous detection model which had begun to produce many false positive detections. The plan was to train a model that is aware of other objects that can commonly be mistaken for a waste bin detector as well as multiple colours of waste bins. After training the model, analysis on its performance showed that our plan worked as intended, resulting in higher accuracy and fewer false positive detections. Once this was confirmed, the model was ready to be deployed to actual waste trucks to begin detecting new waste bins.

The copy paste augmentation pipeline solves a difficult problem. Detection models require large image datasets of the object(s) they are built to detect. Because of this, if an object is very rare or has never been seen in a waste truck, how can we detect this object? This becomes much larger when an object is both rare and dangerous, such as propane tanks and batteries which have led to fires in waste collection facilities. Our solution was to take a small set of images, extract the object, perform multiple transformations to it, and "paste" it into thousands of existing images already in the database.

The Green Screen dashboard is a frontend application providing a quick visual display of waste collection statistics. At a quick glance a user is able to determine how many people in the city, or in a specific municipality, are placing contaminants into their waste bins and whether or not waste trucks are going to need to take their collection to the dump instead of to be recycled.

What Went Well and Why?

Before the fall semester of 2021 even started, we already had a group understanding of our problem and our proposed solution which allowed us to hit the ground running in September. We emphasized good practices from the beginning such as implementing weekly sessions for spring planning and team work sessions. Both of these assured that we would all be aware of what each other were doing so that we could help each other out at any point. We were proud of how organized we were and how well we worked together during our work sessions. We found that allocated work times kept us accountable and helped us keep a consistent pace of completing tasks.

We feel great about having code that we created being deployed onto multiple municipalities trucks. Being able to see our models performance on real world data was extremely satisfying and motivating.

Working towards a software solution to benefit the environment was felt rewarding and fulfilling. The environment is something we all care about and we hope to have a positive impact on it. Creating a solution using a cutting edge method from a research paper was also gratifying. Having no previous applied implementation of copy paste and not knowing whether or not it would be effective was both thrilling and a little terrifying.

The three of us have different backgrounds, experiences, and knowledge bases which greatly benefitted our group problem solving abilities. When tasked with a problem, we found it powerful to have three different perspectives on it and we believe this heightened our ability to find great solutions. This also allows us to allocate tasks to the person who will most efficiently and properly complete them as well as opening up pair-programming opportunities for the others who may not be familiar with a certain technology to learn.

What can be improved and how?

Future Improvements

We initially believed that the copy-paste data augmentation method would be a tool which would require minimal adjustments to fit our needs. This thought process led us to invest an extensive amount of time and resources into conforming this implementation as opposed to taking key concepts and building our own solution. Since this method was new and we lacked core understanding of how it performed, we were hesitant to rebuild this data augmentation method from scratch. However, choosing to do it was the best decision we made as it allowed us to build a custom solution recycling augmentation as well as gave us full control of our project.

If we had chosen this path sooner, our solution could have been more refined. In future projects, we would like to incorporate this problem-solving skill as it builds a deep understanding of how our services performed, but also gives us the control to build a suitable and robust solution which fits our requirements.

Additionally, as our timeline had shifted, we were unable resolve a performance issue we identified with our data augmentation pipeline. We believe AWS after a set period of time was limiting our image upload connection with their S3 bucket. This limit becomes a bottleneck for our data augmentation pipeline. In the future we would like to adequately assess this situation and implement an effective solution.

Advice to Future Groups

First off, teams should not be intimidated by newer methods and/or technologies. With some research, practice, and breaking down your project into tasks it will become very manageable. We also found working with something new to feel gratifying.

We often mention to each other how happy we are that we got started early. Building the required knowledge base for your project and having a plan, even if it is just a rough plan, before even beginning ENSE 400 allowed us to get started very quickly and ultimately achieve a lot more throughout the 8 months of class time.

Allocating work time for planning, working, as well as just chatting helps keep the team organized and on track for completion. We found that having a designated work time to complete non-technical tasks such as scrum presentations and documentation was a great decision. Weekly sprint planning added a subtle pressure of an artificial deadline, although it is obviously okay to take more than a week to accomplish something it was enough to help keep us from procrastinating. Finally, a Discord or Slack server is a must. Discord is where nearly all of our communication took place and is not only a tool for communication but can also serve as a log of notes, meeting notes, files, and credentials.

Lastly, let each other know when they have accomplished something great. Whenever somebody would finish a task, others would often congratulate them, thank them, or let them know that they did a great job. This keeps up the morale of the team and also provides a boost of motivation.

Closing Remarks

All things considered, we found this to be an incredible learning opportunity and an overall wonderful experience. We are grateful and proud at the end of these eight months. We are happy to be delivering on every objective we set out. We feel that we were very ambitious, however, proving to ourselves that we are capable of accomplishing these ambitious ideas through persistence and dedication, is a great boost of confidence. We are glad that we did not shy away from using new ideas and technologies just because they did not have much documentation.

There are many strategies which aided us in our journey which we all are looking to apply in the future. The first of these is to work with a diverse group of people as this helps tackle problems creatively and with multiple perspectives. Working with an industry partner provides a massive benefit of working with people who know more about a subject than you do. They were an outstanding resource whenever we were confused or lost. Another great habit was regular meetings, these keeped us on track and assured us that we were not wasting time. Specifically, sprint planning with a Trello board can also serve as a great log to remember what others are doing and what has been accomplished in the past.

Finally, we wish to mention that we are proud to have worked on a project which improves the sustainability of our society. We hope to encourage other engineers to keep our sustainability in mind with every solution they create.