Worcester Polytechnic Institute

Robotics for Recycling Industry

Authors: Nathan Bargman (CS, HUA), Kaitlyn Fichtner (CS, RBE), Mary Marquette (RBE, ME), Garett Ruping (CS, RBE), Conrad Tulig (CS, RBE), Lauren Wach (RBE)

Advisors: Professor Berk Calli (RBE), Sarah Jane Wodin-Schwartz (ME), Jennifer Rudolph (HUA)

Date: April 28, 2022

Major Qualifying Project (MQP) submitted to the Faculty of the Worcester Polytechnic Institute in partial fulfillment of the requirements for the Degree of Bachelor of Science in Mechanical Engineering, Robotics Engineering, Computer Science and Humanities and Arts.

Table of Contents

Abstract

Our project's goal is to help innovate material recovery facilities' sorting methods required to separate materials before they can be recycled. We designed an autonomous system which is capable of detecting and removing recycled materials from a conveyor belt through the use of deep learning object localization and classification as well as a bi-directional arm with pneumatic suction cups. We also created our own dataset to train our deep learning model and a user interface to correct it during operation. Through these methods we seek to further refine and optimize the way places like Massachusetts sort their waste.

1. Introduction

Countries around the world are moving toward more modern and green initiatives, including more sustainable recycling practices. In the 21st century, reasons to improve these practices continue to pile up. Firstly, the United Nations predicts the human population size is not done growing and likely will not reach a steady number until 2100 [42]. The United States, as an example of the greater problem, is barely keeping up with the waste it produces today and thus has little chance to do so in the future without a significant overhaul to its waste management. Secondly, current poor waste management tactics such as trash piles and indiscriminate incineration threaten the Earth's ecosystems [44], ecosystems that not only animals rely upon but also humans who also need its water and land. Additionally, the consumption and destruction of ecosystems for raw materials such as wood and fossil fuels leads to a growing concern for the sustainability of these practices. Recycling resources like cardboard and paper can reduce the need for logging and fossil fuels by reusing them or converting them into energy. Finally, as the world has gotten better at returning recycled goods back to raw materials, the value of waste has continued to grow. This is to say that the economic reward for recycling waste is starting to eclipse its cost for many categories of trash [36]. All of these reasons together mean that, as mountains of trash continue to pile up all over the world, instituting more efficient and safe recycling practices is more important and rewarding than ever.

Vital to the feasibility of reusing waste, the most important step in recycling is sorting. In order to break down the waste and reuse it, it needs to be as close to pure as possible. For example, a mixed bin of paper, glass and metal cannot be refined, but once each material is sorted into its own pile it can be re-processed into a new good [48]. This means that sorting waste valorizes it. In short, piles of assorted waste are not very valuable, but sorted waste is much more likely to be sold than thrown into a landfill. As such, separating trash and picking out valuable materials is a straightforward way to make money.

For many years companies in countries like China would buy overseas trash to sort through and then sell the resulting resources for a profit. Labor in China was cheap and the sheer amount of trash available for purchase meant that there was a lot of money in sorting trash, so much so that civilians would illegally traverse massive, extremely dangerous trash piles to pick out materials to sell [45]. This method of recycling was far from perfect; it did, though, remove a lot of the sorting responsibility from the West and placed it onto China. Cities like Boston in Massachusetts specifically, exported nearly 90% of their waste to China [35]. Recently, however, China stopped buying trash in order to deal with its own waste, reduce detrimental environmental effects to its nature, and further its own political objective of being less reliant on other countries. Exporting waste to China was good for both countries as foriegn waste could be cheaply refined and used to fuel much of China's massive industrial sector. Although other smaller countries continue to buy recyclable waste, with a majority buyer gone, many countries like the United States were left with heaps of unsorted garbage [45]. Without a way to sort it, they just have piles of recyclable materials that cannot be recycled. This need for better and more efficient sorting methods is the impetus for our project.

Our project is the design and construction of a robotic system to pick out specific types of trash on a conveyor belt. In context we first inspected a materials recovery facility (MRF) which currently has several manual and automated sorting stages. We explored how a robotic system can be designed to help this sorting process. Our robotic system employs Deep Learning to sense objects on the conveyor belt and uses a vacuum system to pick them up as needed. However, because the arm and associated software will not be 100 percent successful in a fully autonomous mode, the system is also designed to take manual inputs from a worker using a human machine interface for correcting the machine when it can't identify or misidentifies objects.

Cardboard, a commonly recycled material, is difficult to sort due to its high variability. Cardboard is used in such a wide range of commodities requiring different shapes, sizes and material makeups. This means the construction of one system to sort all of these types is difficult. Current systems in cities like Worcester employ a type of screen that shoves large items in one direction while smaller items fall through [48]. While this is effective at filtering out large pieces, small pieces of cardboard remain unfiltered. To muddle the issue even further wet cardboard and torn or destroyed cardboard are extremely difficult to filter out as they can even jam systems or squeeze through small holes designed for other materials. Building upon previous year's work on this project we decided to narrow focus on this problem.

2. Background

Providing automation solutions in any sector can be costly; it requires significant investments from the companies and can significantly affect jobs in the industry. We want to ensure that our project has a positive effect on the industry and waste sorting in general. This means that we will not only consider the ethical implications of our design but the design itself. In other words, to affirm that our design retains human interaction is insufficient when deciding if our project productively modifies the current state of technology in the field. To situate this project in that field this report will explore a number of modern recycling waste management solutions around the world. We analyze, compare and contrast the systems of Worcester, Massachusetts (the home city of this project); Pordenone Province, Italy; YanYuan, China; and Kamikatsu, Japan to determine best practice for Worcester.

2.1 Best Practices

Worcester, Massachusetts

For its municipal recycling Worcester, located in central Massachusetts, contracts the company Casella, who boasts the single-stream (zero-sort) recycling system. In single-stream recycling the households each put all of their recycling in their own single bin, all of which are then brought to a MRF [48]. Typically recycling is sorted using a combination of different automated and manual systems. As the mixed recycling arrives at the MRF, it is pushed onto a conveyor belt where extremely large items and non-recyclable items are manually removed. Then the materials run over an 8-inch screen, where large pieces of cardboard are automatically removed from the feed while smaller objects fall through. Then the fallen materials are pushed through a glass breaking screen which smashes glass into small pieces that fall through into a collection bin where the glass is sifted before removal. The rest of the materials move under a large magnet removing any steel from the recycling. Paper is then removed with a steep angled screen while the rest of the materials, mostly plastic and aluminum, are separated using an optical sorter and eddy current (i.e. light refraction and electric currents). After the separation process, the sorted waste is then sold to companies who can make use of the raw materials. This intensive process is complicated, and it effectively removes almost all sorting responsibility from Worcester residents.

This system's biggest flaw is that it can’t handle certain common items such as plastic bags, any sort of rope-like material or any item smaller than two inches square. These items can either clog the machines or bypass them entirely. As previously mentioned in order to properly recycle materials they must be as close to pure as possible. This means that if materials get sorted into incorrect piles they can completely ruin the whole process, rendering the materials unrecyclable. In addition, these incompatible items tend to create safety concerns inside the MRF. For instance, when the sorting screens become clogged workers have to carefully enter the machine and cut away the offending materials [48]. Thus although this system removes most civilian responsibility, it is still reliant on people following some guidelines for recycling.

Worcester recycles 32% of its municipal waste [47]. In order to try to improve this metric the Massachusetts state government and Casella attempt to provide easy access to educational information online on what kinds of waste should be recycled and what should not [34].

Pordenone Province, Italy

Pordenone Province, located in northeast Italy, requires residents to sort their own recyclable waste. Plastic, paper, cardboard and metals must be placed into different bins before they are transferred toa MRF. These MRFs are different from those found near Worcester, as they perform a secondary sort and do not have to sort all of the recycling. Waste is transferred through a variety of methods, depending on the community. Some communities even have specially designed bins that record weight for better tracking of waste. In addition to individual home sorting, communities also have common collection points located on street corners and sidewalks [36]. Since this system is relatively simple, its biggest drawback is that it is very reliant on residents' knowledge and compliance. To encourage proper compliance, if residents or even tourists do not properly sort their waste, they will incur fines or legal action [36].

Pordenone recycles a total of 41% of their waste, and Italy in general is among the best recyclers in Europe, with a greater than 50% recycling rate [36]. This difference is likely because each community in Italy employs different regiments, including collection and enforcement methods. As MRF facilities are less important, much of the focus on improving recycling lies with companies who buy, sort and refine the raw materials. For example Green Plast, an Italian based plastic and rubber recycling initiative, holds an annual conference showing off new refining equipment as well as best practice for handling those materials [40].

Yanyuan, China

Situated in Sichuan Province, Yuanyuan county is home to several small cities in western China. To recycle waste, residents must sort waste into separate bins. These bins are communal and emptied informally by individual workers employed by the state [45]. Actual facilities only facilitate less than 10% of the recycling processed in Yuanyuan. This practice is derived from the waste picking work that happens in tandem to ongoing recycling. For many years the extreme amount of waste China produced and bought would often end up in enormous trash piles. There, trash pickers would traverse across the mountains of trash picking out any material of value [45]. To help phase out this hazardous work, civilians who would typically do trash picking are given the opportunity by the state to act as material curriers between rural communities and places where resources are stored or processed. This once informal process has become an encouraged job to allow resources from remote areas to be recycled and reused.

Outside of Yanyuan, in places of higher technology, the communal bins are integrated with a reward system for those who participate. Residents may have to scan their face before they deposit their waste and those who do regularly and correctly are rewarded monetarily or in everyday goods. This type of system has helped the city of Changsha achieve a 100% participation rate and 70% proper sorting rate [41].

Sichuan recycles around 34% of its waste [32], while China as a whole currently only recycles around 20% of its waste [41]. To remedy this, China is trying to improve its waste management by designing Zero-waste cities that will be breeding grounds for innovation in recycling under strict governmental supervision [33]. In addition, since 2009 China has become more aware of how much solid waste it produces choosing to incinerate nearly 50% of its municipal solid waste from major cities in 2018 [44]. With increased efforts to match heighted demand for better management systems, China is setting itself up to be a world leader in reducing solid waste, but its recent changes have yet to be implemented nationally and cause significant improvement. To hasten this transition China has been encouraging the education of small towns [45] and continues to build major infrastructure throughout the country. As recent as 2018 China completed major construction of roads to eastern China and Tibet as part of their Belt and Road initiative [46]. These roads allow new commerce and trading which would include resources from recycling.

Kamikatsu, Japan

A small town in Tokushima Japan, Kamikatsu, is home to around 1400 people and is trying to revolutionize waste management [50]. Per its own initiative Kamikatsu has been aiming to recycle one hundred percent of its waste for almost two decades. In order to achieve this goal, the residents of the town separate their waste into forty-five different categories. The categories include but are not limited to different types of metal, plastic, glass and paper. Once separated, waste is delivered to a Kamikatsu's Zero Waste Center where materials are reused, composted, re-sold or incinerated. Almost all of the waste is re-purposed, leaving less than 20% of waste in landfills or incinerators [43]. This herculean effort is only possible through individual motivation as each resident must manually deliver and sort their own waste. This attitude is highlighted in the construction of the collection point which is designed to look like a big question mark and given the provocative moniker of WHY [51]. Kamikatsu shows just how much can be reused given a dedication and drive to make a difference. However, despite their best efforts, according to Kamikatsu's chief environmental officer, materials like diapers and heat packs are very difficult to recycle [43]. As such reaching 100% waste reduction is near impossible to obtain without improvement to today's technology.

Zooming out, Japan in its entirety recycles around 20% of its total waste [38]; however, it excels in plastic recycling with approximately 85% of Japan's plastic waste being collected to be recycled [39]. This large number is made even more impressive when considering that Japan as a nation uses an incredible amount of single-use plastics. Even products including produce are all wrapped in single-use plastics to accommodate Japan's working lifestyle. Through many years of government-led academic initiatives, Japan was able to dramatically increase its plastic recycling rates. This education is evident in the reverence and fascination Japanese citizens have for recycling [38]. This achievement however is dwarfed by a faulty backend. Because Japan used to ship the majority of its plastic to China, infrastructure had never been set in place to actually process the recycled plastic. As a result, now Japan has no choice but to burn the plastic it has no way to get rid of [39]. While incineration can lead to energy generation, as mentioned previously incineration is detrimental to the environment. To fix this issue, Japan is focusing its efforts on advancing the current technology available to break down and repurpose plastics [37, 39]. Like with Kamikatsu, beyond increasing the percentage of waste recycled, innovation to ensure that material can actually be reused is in turn just as important.

2.2 Analysis & Considerations

In analyzing different sorting practices two main methods emerge, residential sorting and centralized sorting. Residential sorting requires residents to sort their own waste and either deposit that waste in a centralized location or prepare that waste to be collected near their residence. Centralized sorting means waste is collected en masse and sorted in a centralized location before being processed separately.

2.2.1 Residential Sorting

In the four cities analyzed, residential sorting emerged as the most efficient (i.e., having the greatest recycling rate) and therefore logically the most popular. All of the locations analyzed had some form of residential sorting. Worcester, too, uses this strategy to sort items that cannot be sorted in its centralized sorting system [48]. This method's efficiency can be attributed to several different aspects. Primarily it introduces a social aspect into recycling. Forcing residents to either dispose or sort their waste in a public place creates a source of peer pressure to follow the rules and look good in front of others. A disregard for the rules would likely end in social embarrassment or even becoming an outcast. As people tend to avoid punishments; locations that use residential sorting exclusively enjoy the highest efficiency. Moreover in places where recycling has become a social norm, like Japan, this sort of peer pressure helps enforce and maintain such practices [38]. The strongest example of this among the case studies here is Kamikatsu. The only method for residents of Kamikatsu to dispose of their waste is to sort it at a community collection point. Residents are more than willing to dedicate a serious amount of time to sort their waste into a few dozen categories. This is at least in part because social pressure is maximized in Kamikatsu as it is a very small community where recycling is very closely monitored by everyone. An individual not following procedure would be much easier to notice in Kamikatsu than in larger cities like those in Italy and China. For this reason, Pordenone

Italy takes advantage of another benefit of residential sorting. Since the collection point can be in a public space it would be easy to institute a surveillance system or modify the collection bins in some way to help enforce procedure. Monitoring collection means that cities can enforce compliance with the threat of punishment. Or conversely Changsha China, rewards compliance by closely watching which residents sort their waste correctly and punishes those who don’t. This type of enforcement doesn't just affect Italian citizens though, tourists create a non-insignificant amount of waste and a public collection point ensures that tourists can’t help but notice it and either learn by example or curiosity [36]. These motivations are vital to the efficiency of residential sorting. Since MRFs in places with residential sorting require at least partially sorted waste to function, motivation for civilians to take heightened responsibility for their waste is required.

Individuals that have an altruistic sense to save the environment or live in a recycling culture may be more likely to recycle, but if there is not a rigorous system in place to ensure correctness efficiency will suffer. As seen in Changsha, although 100% residential sorting participation is possible, making residents correctly sort their trash is difficult even with a surveillance system. As such to enhance the system further residents need to be educated on how to properly dispose of waste. This is further supported by China’s focus on education in places like Yuanyan as they are confident in their ability to achieve high participation rates through governmental supervision. Education, if done correctly, can even eliminate the need for extreme supervision. Although Japanese citizens have to meticulously sort their plastic into many different categories, Japan has the highest recycling rate for plastic in the world at over 85%. This efficiency did not start overnight however, as Japan made serious investments into educating their citizens on how to recycle. In the 1980s and 1990s, Japan led many initiatives via mass media, private businesses and government operations which led them to create a generational change [38]. This investment paid off, as today Japan has not only built a culture that values recycling, but also one that understands how to recycle. Through the popularity of the Zero-waste concept in Japan it is evident that the education the population received had a lasting impact [51]. Education that creates individuals who are self motivated and trained can dramatically increase a country's efficiency. Replicating Japan’s results would be very difficult to do in the short term as it took Japan 30-odd years to see major benefits from their initiative. China, for example, may see benefits in major cities, but since their infrastructure is still growing, expanding to a nationwide effort would be tricky. The United States, furthermore, would have difficulty gaining traction among private business and media across the whole country.

In all, residential sorting seems to be an obvious choice when looking at the results of cities like Kamikatsu. However, Kamikatsu has an extremely low population size and an unusually high motivation for recycling. China’s recycling rates are rising, but China is in a unique position as they have a relatively high amount of governmental power. This means that the government can demand provinces which in turn can demand cities to reach certain goals. If they fall short, they will face the consequences served by their powerful government. As such residents in cities where this strategy is deployed, are highly motivated by law to follow the rules. This type of strict policy would not work in every country, as many governments do not have the power to demand such a task with the threat of a severe punishment. Italy sits in the middle of the two aforementioned models, not having as much governmental power as China and not as much social motivation for recycling as Kamikatsu. Despite that, a 50% recycling rate country wide is testament to how effective the Pordenone method is. Either directly enforcing participation and correctness or indirectly through education and culture causes high efficiency in residential sorting.

2.2.2 Centralized Sorting

A centralized sorting system removes as much responsibility from individuals as is currently possible. Residents are only required to place their non-food solid waste into one bin and exclude items dictated by their local MRF. Therefore, a centralized sorting system relies less on law enforcement and education and instead is more focused on how the MRF itself works. A drawback of this approach is a city's recycling will only ever be as efficient as its MRF. In other words the MRF acts as the bottleneck as opposed to residential sorting where participation or education are often the limiting factors. Moreover, while residential sorting entails lots of individuals sorting small amounts of waste, centralized sorting requires one location to sort massive amounts of waste very quickly. This is a difficult task carried out by a mix of autonomous and manual processes. There are inaccuracies and inefficiencies in this sorting process and injuries can happen often due to sharp or unwieldy objects to avoid damaging the machinery. Even small issues that take minutes to fix can leave metric tons of trash unprocessed. These efficiency ceilings explain why Worcester's current MRF-based recycling might not be operating as efficiently as cities like Kamikatsu.

Centralized sorting allows the MRF to easily dictate where unsortable or unrecoverable waste ends up. In Worcester much of that waste ends up in a landfill where its gasses are harvested for power generation. In contrast in other countries like Italy, companies may have to give serious investment to pre-sorting the waste they buy so that the materials don't become unusable [40]. As mentioned previously if a material isn't pure, it can completely ruin the whole recycling process. A centralized sorting system eliminates this step and eases the demand on recycling plants to prepare their materials before recycling.

Improving a centralized sorting system requires residents to sort their materials in accordance with what the MRF prescribes. Another way to proceed, however, is to improve the MRF itself to accommodate a wider array of materials. If that is achieved, then resident responsibility could be entirely eliminated. This kind of change happened in Seattle, Washington, resulting in its system reaching a high rate of 56.9% efficiency [49]. Cities like Worcester could learn from the logistic and facility changes that Seattle made to try to replicate their results without having to change legislative policy or involve every resident in the area. Places like Japan and Italy however would require large amounts of infrastructure change to build and maintain MRFs capable of sorting all waste at once. China on the other hand has plenty of labor and money to institute this kind of system. However it would contradict much of the work and culture that was built up from the informal sorting of waste for many years.

2.3 Synopsis

Countries around the world are moving towards greener initiatives and technology. It is important to understand which methods will have a positive impact and which will have a negative one. As developers of such technology thoroughly analyzing our project beyond just its immediate impact will ensure that it is a lasting positive force. Worcester's current centralized recycling system fits the city versus a residential sorting system. Worcester lacks the ability to sufficiently motivate citizens to sort their own waste. That is to say that it would not be possible to institute surveillance systems and it is unreasonable to expect a sudden shift in culture or education. Furthermore Worcester's use of MRFs means that education is less integral to the system's efficiency. In fact projects like ours begin to contribute to more relaxed regulations for what items MRFs can handle. Although residential sorting systems show the greatest efficiency around the world, centralized systems with help from said projects are capable of reaching high recycling percentages in the future. In all our project, built for a centralized sorting system, is a building block for future projects at WPI and beyond in cities like Worcester.

3. Analysis of previous year’s work

This project was started by an MQP team in the 2019-2020 academic year and was continued by another MQP team throughout the 2020-2021 academic year. The previous teams designed a robot frame, an arm, a gripper, and a conveyor belt that sits below the robot. This section describes the state of the waste sorting system at the beginning of the 2021-2022 academic year and outlines our goals to improve the robot’s ability to sort recycled waste.

3.1 Mechanical

3.1.1 Robot

The robot itself was originally designed and assembled by the first few iterations of this project. The overall robot is held together by a welded and bolt-together frame. The robot moves along steel X- and Y-axis rails; the direction of the axes in regard to the robot can be seen in Figure 3.1. Moving the robot along the rails by hand the team noticed that there was not a smooth movement and that loose wires connected to the robot were getting stuck during motion. Two carts from the X-axis of the robot sit on top of the two carts on the Y-axis. The carts are connected by a pair of linear rails. The X-axis belt is controlled by one of the carts and the other cart is used to support the pulley for the belt. As can be seen in Figure 3.2 when initially evaluating the system it was noticed that the cart plates where the linear rails attached to did not sit parallel on top of each other, which caused awkward movement of the carts and the robot along each of the axes.

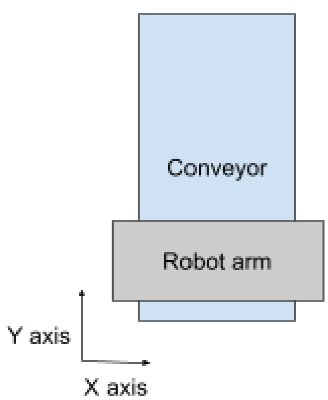


Figure 3.1: Robot Coordinate Frame   Figure 3.2: Upper cart sitting unevenly on the lower cart

The following solutions were identified to help improve movement of the robot:

1. Grease the rails on the X- and Y- axis
2. Fix loose wires by running them through enclosure to ensure robot safety
3. Adjust the cart plates so they sit parallel on top of each other

3.1.2 Conveyor Belt

The previous team was able to design and build a conveyor belt for testing the robot. The belt is reversible, able to operate at varying speeds, and can travel up to 60 feet per minute. Manual adjustment of the belt was done through tightening the tensioners on the ends. There are still some problems with the belt slipping, but the team did not find this problem to be a priority for the time being.

Figure 3.3: Conveyor Belt

3.1.3 Gripper

The gripper created by the previous team was designed with spikes so that it would be able to latch onto a piece of waste. It contained a central rotating component controlled by a Dynamixel motor to allow the gripper to fully close which created a pinching action to work in combination with the piercers. Upon further evaluation of the gripper a few different problems were identified. First, the gripper does not reach the conveyor belt because the arm is not currently long enough to reach. The next problem was that the end effector was not rigid. This meant that oftentimes when the gripper came into contact with a piece of recycling it would swing to the side the actual gripping mechanism was unable to interface with the recycling. In Figure 3.4 the gripper is being pushed to the side because it rotates on a bolt that connects it to the arm. Along with the previous team's design not being rigid it also only had the ability to translate along the X- and Y-axes meaning it was unable to pierce or grasp flat pieces of cardboard.

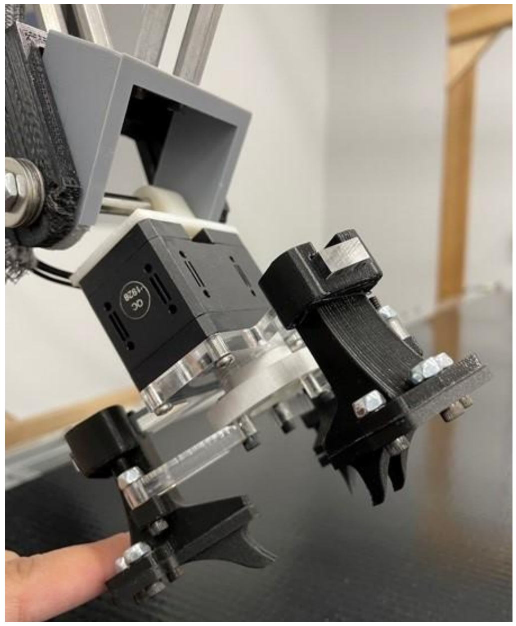


Figure 3.4: Previous gripper being pushed to the side

Through testing the team discovered that the piercers on the end effector did not support picking up many different kinds of cardboard including wet cardboard, boxes, crumpled cardboard, thinner paper cardboard, etc. If the end effector was moving it often was not strong enough to hold a piece of recycling without dropping it.

The solutions below were identified as ways to help improve the gripper:

1. Increase the length of the gripper arm or size of the gripper to reach the conveyor belt
2. Make the end effector semi-rigid to allow it to interface with a piece of recycling while still having the ability to adjust slightly if needed
3. Redesign the gripper to better grasp dry, wet, crumpled, and flat cardboard

3.2 Software

3.2.1 Robot Control

The codebase that last year's MQP provided had a working GUI that allowed for moving the robot arm along the X axis, a homing sequence for the arm to start in the origin, a control system for closing and moving the gripper up and down, and a camera feed from an Intel RealSense Depth Camera D435. The original system for moving the robot used a user-provided amount of steps for the stepper motor to move the arm, which made movement difficult, requiring converting distance to stepper steps. The code also had no system to move the robot arm along the Y axis which was identified as possibly being useful to implement. Similarly the homing system did not move in the Y axis. Although control of the gripper system worked, it would need to be changed for future gripper designs. The camera feed converted the data into a cv2 image which could be used to identify trash objects with computer vision.

The existing code had no logic for an automated system for picking up trash items. The team decided that new Robot Operating System (ROS) nodes would be necessary. One node that takes in images and identifies the position of recycling items, and another node that would move the robot to the position of the recycling item, wait for it to be underneath the arm, and pick the item up. A new GUI would also be needed to show the recycling identification procedure.

The following solutions were identified to help improve robot control:

1. Convert stepper step to centimeters when specifying arm position
2. Add Y axis movement for the arm, and add that to the homing procedure
3. Add a camera node and calculate camera transformations to integrate live object detection
4. Adjust gripper control for a new gripper design

3.2.2 Computer Vision

Minimal progress was made in computer vision and detecting cardboard in previous iterations of this project. Last year's team trained a YOLOv3 [13] model on a dataset of 85 images selected from TrashNet [1]. They also increased their dataset through data augmentation with horizontal and vertical mirrors and 90, 180, and 270 degree rotations. The YOLOv3 [13] model was tested on 12 images taken in the lab. While specific mAP and accuracy scores were not provided, the report mentioned that one piece of cardboard was identified in 2 of the 12 test images.

Upon examining the TrashNet [1] dataset, we found that the images were not representative of a real MRF recycling stream. In addition, the YOLOv3 [13] model is outdated now with better state-of-the-art models being released since then. It is decided that the computer vision system needs to be redesigned and implemented from the ground-up.

4. Methodology

This section describes the research and rationale behind our approach to improving the robot's gripping capabilities and object detection algorithm. We also introduce a sorting procedure, near-IR research, and a user interface for human-robot interaction.

4.1 Mechanical

4.1.1 Gripper

The original gripper mechanism was designed to combine a pinching action with the piercing action of the prongs located at the tips of the gripper. Although the previous year's team had modified this gripper mechanism to be able to more effectively puncture materials, our team noticed that the design was not strong enough to be able to puncture cardboard or pick up pieces that laid flat on the conveyor belt. The mechanical aspect our team wanted to focus on was being able to have a gripper that would be able to pick up cardboard more reliably. In order to be able to implement a more improved gripping mechanism we researched different types of gripping mechanisms. The three types of mechanisms that we looked at were mechanical, amorphous, and vacuum.

4.1.1 Mechanical

The first type of gripper that the team researched was a mechanical gripper because it would be the easiest to integrate into the current system. Mechanical grippers convert drive motion into gripping motion of the gripper jaws. There are two main types of actuation used in mechanical grippers: pneumatics and electrical motors. This actuation is used to operate the jaws or "fingers" that are actually grabbing the object. Configurations of the mechanical grippers usually have either a 2-finger or 3-finger arrangement [26].

There are many different types of mechanical grippers that each excel at picking up different materials. The two types of grippers we focused our research on were the parallel motion two-jaw gripper and the three jaw gripper. These grippers usually excel at picking up rigid three-dimensional objects. Based on initial testing of the previous gripper and research we learned that they are not ideal when picking up flat items. Considering our project goal, which was separating mostly flat cardboard items from the waste system, we decided not to pursue the mechanical gripper route.

4.1.2 Amorphous

The team also looked into amorphous grippers. An amorphous gripper is able to mold itself to any object. It uses a process called "jamming." When the granular material inside the end effector is compressed the gripper becomes very rigid. This allows the griper to deform around its target while it is still in its malleable configuration, and then harden around the object it is trying to pick up when jamming is initiated [27]. This is depicted in Figure 4.1.

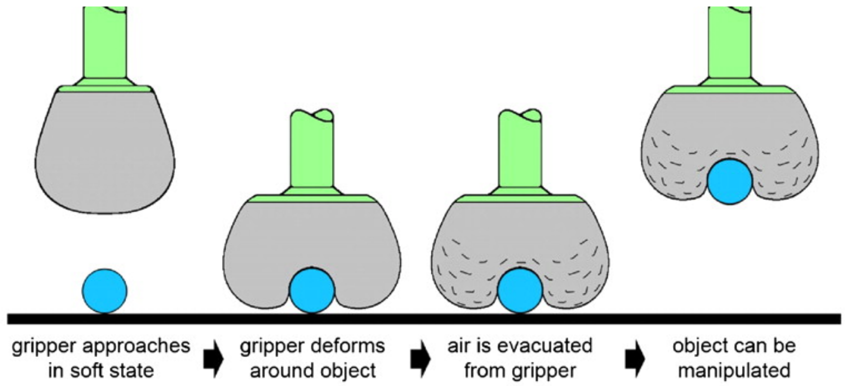


Figure 4.1: Depiction of Amorphous Gripping [30]

The gripper strength is seen to be due to three mechanisms which are all controlled by jamming [27]. These mechanisms include interlocking between the gripper and object surfaces, which causes geometric constraints. The jamming also causes static friction from the normal stresses at contact with an object and an additional suction effect that can occur if part of the object's surface is sealed off by the gripper membrane [27].

In order to test this design preliminarily for ourselves, a basic model of an amorphous gripper was constructed by the team, as seen in Figure 4.2. A balloon filled with coffee grounds served as our end effector which was then attached to a funnel to connect to our vacuum component. We used a balloon because we knew that the gripper had to be malleable without vacuum and coffee grounds were a granular enough material for our uses. Once the vacuum was turned on, the coffee grounds were compressed and the balloon became rigid. While the gripper was successful at picking medium sized three dimensional objects, it mostly failed for very small and flat objects. We realized that this type of gripper would be unable to work for our purposes and continued our research of other mechanisms.

Figure 4.2: Mockup of Amorphous Gripper

4.1.3 Vacuum

The last type of gripping mechanism we looked into was vacuum grippers. Vacuum grippers use the difference between atmospheric pressure and vacuum to lift, hold, and move objects. The vacuum is generated by a compressed air-driven pump, such as a compressor, or an electromechanical pump. Compressors are able to produce up to ten times more power than electromechanical pumps [28].

Vacuums have the ability to pick up many different types of materials as long as there is a tight seal between the object that is being picked up and the suction cups. This means that the vacuum gripper would have the ability to pick up flat objects along with the other objects, like cubes or boxes. They can also pick items up in most positions as long as the surface they are trying to connect to has a larger area than the diameter of the suction cup being used [28]. The drawbacks to vacuum systems are that they are more expensive than the other options discussed and would struggle in dusty conditions. Having a lot of dust in the air near the vacuum system would mean that the system would need to have filters to keep the dust out of the vacuum and compressor. In dusty conditions the suction cups would also need to be cleaned or replaced frequently.

A vacuum gripping mechanism was ultimately decided upon after our research into different types of grippers. It would be able to pick up flat materials such as the cardboard we wanted to work with for the purposes of our project. The materials would not be too expensive for our work and the mechanism seemed to be the most reliable option.

4.2 Software

4.2.1 Object Detection

For the robot to detect, localize and classify recycling on the conveyor belt, we needed to train a deep learning model. Because recycling streams are extremely cluttered, detecting recycled materials in a real waste sorting setting is a challenging task.

Dataset

To train an object detection model, we needed to find a dataset that demonstrates the highly cluttered and occluded environment of a recycling stream. Multiple open-source datasets exist in the recycling domain, including TrashNet [1] used by the previous year's team, Trash Annotation in Context (TACO) [2], and ReSortIT [3]. TrashNet consists of 2,527 images with one annotation per image across six classes of trash. Because each TrashNet image contains only one piece of trash with a solid background, we determined that the dataset would be too simplistic and not allow the model to learn the complexities of a real recycling stream setting. TACO consists of 1,500 annotated images with an average of 3.2 annotations per image across 60 classes of litter. Although TACO images contain multiple pieces of trash, they still have a low level of clutter compared to a recycling stream. Additionally, TACO images have mostly outdoor backgrounds which look very different from a material recovery conveyor belt. ReSortIT's Synthetic Complex dataset consists of masked recycling objects randomly placed in pairs against colorful backgrounds. ReSortIT uses 1,600 images of individual recycling objects across four classes to create a synthetic dataset of 21,600 annotated images. ReSortIT is a promising dataset because the randomness of the backgrounds is useful for ensuring the model learns the recycling type independent of foreground objects and background. However, ReSortIT has only 2 objects per image which is a low level of clutter and many objects appear distorted due to the masking process of creating the synthetic dataset.

ZeroWaste [4] is a dataset that uses video footage of the waste sorting stream at a real Materials Recovery Facilities (MRF). ZeroWaste's fully annotated dataset, ZeroWaste-f, consists of 4,503 images with an average of 6.1 annotations per image. Another version of the fully annotated ZeroWaste dataset, ZeroWasteAug, adds augmented objects to the rarer classes in ZeroWaste-f using objects cropped out of the TACO [2] dataset. Including both the real and augmented objects, ZeroWasteAug has a total of 38,328 annotations with an average of 8.5 annotations per image. ZeroWaste has a large number of annotations, a high level of clutter and occlusion, and a real MRF conveyor belt background. We determined ZeroWasteAug to be the best dataset to train our model on because it has the most realistic setting and clutter level compared to other recycling datasets. It also has a more balanced class distribution compared to ZeroWaste-f.

Table 4.1: Evaluation of existing recycling datasets [1, 2, 3, 4] by domain, number of recycling classes represented, total number of images, total number of annotations, and clutter level of each dataset. Clutter level is measured as the average number of annotations per image.

	Domain	# Classes	# Images	# Annotations	Clutter
TrashNet	Real	6	2,527	2,527	1.0
TACO	Real	60	1,500	4,784	3.2
ReSortIt	Synthetic	4	21,600	43,200	2.0
ZeroWaste-f	Real	4	4,503	27,744	6.1
ZeroWasteAug	Real	4	4,503	38,328	8.5

Figure 4.3: Sample annotated images from ZeroWasteAug

Although ZeroWasteAug is a strong dataset due to its realistic setting and the large quantity of annotations, such a high level of clutter and occlusions, as seen in Figure 4.3, makes it challenging for the model to learn. ZeroWaste [4] trained RetinaNet [6], Mask R-CNN [7], and TridentNet [8] models using weights pretrained on the large-scale MS-COCO [5] dataset and further fine-tuning on the ZeroWaste-f dataset. In ZeroWaste's experiments, TridentNet performed the best of these three models with 24.2 mean average precision (AP). These results indicate how challenging this domain is for state-of-the-art detection models.

Since we would expect to have domain adaptation challenges if we trained our detection model on only the ZeroWaste dataset, we decided to create our own dataset on a conveyor belt with approximately 50% the clutter of ZeroWasteAug. Creating a dataset at our lab would improve our ability to detect recycling on our lab conveyor belt, making it easier to run experiments and test our sorting mechanism. Moreover, our dataset could be used in combination with existing recycling datasets to potentially improve object-detection results in this domain. The design of our dataset is described further in Section 5.2.2.

Model & Training

The following section describes our research on deep learning models to implement for our recycling objects detection task. Since ZeroWaste [4] is the dataset we selected for training our model, we began by looking into several models ZeroWaste tested on their dataset. ZeroWaste's best results used the semantic segmentation model, DeepLabv3+ [9], achieving 52.23 mIoU and 84.79 pixel accuracy. ZeroWaste also trained popular models RetinaNet [6], Mask R-CNN [7], and TridentNet [8] with resulting mAP scores of 21.0, 22.8, and 24.2 respectively. These results are promising, but demonstrate the challenging nature of detection in this domain.

Because our task of picking up recycling objects requires only locating the center of each object, we decided that bounding box predictions would be sufficient for our application. Predicting bounding boxes would reduce the annotation workload in the creation of our own dataset, simplify calculations used in our sorting procedure, and allow us to use smaller models with faster inference times.

Since utilizing the masking portion of DeepLabv3+ [9] and Mask R-CNN [7] was not within the scope of our project and to increase the diversity of researched models in the field, we chose to experiment with Faster R-CNN [12], a near state-of-the-art predecessor to Mask R-CNN. While Faster-RCNN is a large model at ~800 MB, it is still a successful and widely used deep learning model. Faster R-CNN is a standard two-stage convolutional neural network (CNN) model where the initial CNN layers in the first stage are then used in two separate networks: one for classification and one for bounding box localization.

We also wanted to implement a second model that focused on small model size and fast runtime. The YOLOv3 [13] model implemented by last year's team is a recent near state-of-the-art model that achieves high accuracies and low runtimes due to YOLO's unique architecture. YOLO (You Only Look Once) is a one-stage object detection method, meaning that both classification and localization happen together. This enables YOLO's very fast runtimes. However, since YOLOv3 has been released, an entire family of YOLOv5 [14] models have been developed that provide differing sized models achieving either faster runtimes or higher accuracies than YOLOv3. We also researched additional papers [15, 16] that classified recycling in Trashnet and found promising results with ResNet [17], Inception [18], and Mobilenet [19] models. While ResNet and Inception models are already similar in size to a Faster-RCNN [12] model, Mobilenet is another fast runtime model to consider. Based on findings directly comparing Mobilenet to various YOLO models [20], the smallest YOLOv5 model was found to yield both higher accuracy and lower runtime over Mobilenet. As a result, we chose to implement YOLOv5s as our second model with a size of 14 MB.

4.2.2 Object Tracking

Utilizing the trained object detection model, we considered a few systems for tracking the detected objects as they traveled down the belt each frame. One solution we considered was to train a recurrent neural network (RNN) where the classifications of the previous frame help influence the classifications being made in the current frame. However, we determined a couple of roadblocks with this option. RNNs require a large training set as they need many videos to train on [21]. With the time and resources available to us we were only able to record and annotate a half dozen videos. Additionally, RNNs are generally more resource intensive during runtime compared to a small CNN like YOLOv5s [22]. Since our lab computer lacked a dedicated GPU, we opted to prioritize generating a mathematical solution in conjunction with our YOLO model.

For our mathematical solution, we developed a system that continuously finds, predicts and updates the locations of each recycling object. To do this, we took advantage of the fact that all movement between frames is a constant rate and direction on the conveyor belt. As such, we account for the moving conveyor belt by shifting the bounding box's position linearly each frame. Upon receiving a new frame with newly classified bounding boxes, we are also able to compare new and old bounding boxes by taking an intersection over union (IoU) between the new box and the old boxes shifted into a predicted position. Bounding boxes with high enough overlap are then determined to be the same object from the previous frame. Lastly, we keep track of the object's class each frame and use a kalman filter across all frames to find the object's highest confidence class. This process is explained in more detail in Section 5.2.3.

4.2.3 Human Robot Interaction

Automation of the recycling sorting process is a highly complex and difficult computer vision task that could benefit greatly from human interaction. Studies show that human-robot collaboration can significantly improve performance and productivity in the workplace. For example, a survey of 1,075 companies across 12 different industries found that companies implementing non-collaborative automation improved performance by a factor of 2 while companies adopting all 5 principles of human-machine collaboration improved performance by a factor of about 6.5 [11]. In the task of robotic waste sorting, human workers could act as quality control to ensure the robot correctly locates and classifies pieces of recycling. As suggested in ZeroWaste's research, a human-robot workflow for waste sorting could maximize efficiency, profit, safety and work quality in MRFs [4].

An interactive user interface for human-aided recycling detection could be helpful on multiple fronts. Current state-of-the-art detection methods have yet to reach high accuracies on cluttered recycling streams, so human corrections of computer-predicted labels could greatly improve performance. Human corrections can provide more training instances for the detection models to learn on while also giving live feedback to the robot on what pieces of recycling to pick up. Furthermore, allowing MRF workers to sort through a user interface instead of by hand could improve safety and working conditions at material recovery facilities.

4.2.4 Near IR

Analyzing light spectrums to assist in the classification of waste materials is not a new concept. Local MRFs to our project employ optical sorting to sort types of aluminum, paper and plastic. Specifically a camera scans each object before sending it in a different direction depending on what frequencies of light that camera receives. This method does not use deep learning and merely observes one object at a time and makes a simple prediction based on the light it reflects [48]. Current deep learning literature uses this type of analysis to aid in the classification of organic materials, but to our knowledge does not explore how materials such as waste could be better classified when adding additional spectral analysis [24].

To explore this, we decided to research the feasibility of adding an additional spectrum to our object detection setup. We chose NIR over other spectrums such as IR or UV, because it is used in optical waste sorting [25] and in literature which incorporates similar deep learning strategies to our own [24]. If types of waste are differentiable on the NIR spectrum, as the existence of optical sorting implies, then using it as an additional dimension in our deep learning analysis would allow for greater accuracy in the classification of detected objects.

5. Design & Implementation

5.1 Mechanical

5.1.1 Robot

Wire Harness & Electrical Enclosure

During our initial overview of the robot we noticed loose wires that were not plugged into the limit switches that help align the x-axis of the robot. To fix this issue, we added wire connectors to the wires and plugged them into power and ground on the limit switch. These wires were also dangling creating a safety hazard when the gripper was traversing across the linear rails (see Figure 5.1). To eliminate this hazard, we added the limit switch wires to a wire harness that runs along the frame of the robot as can be seen in Figure 5.2.

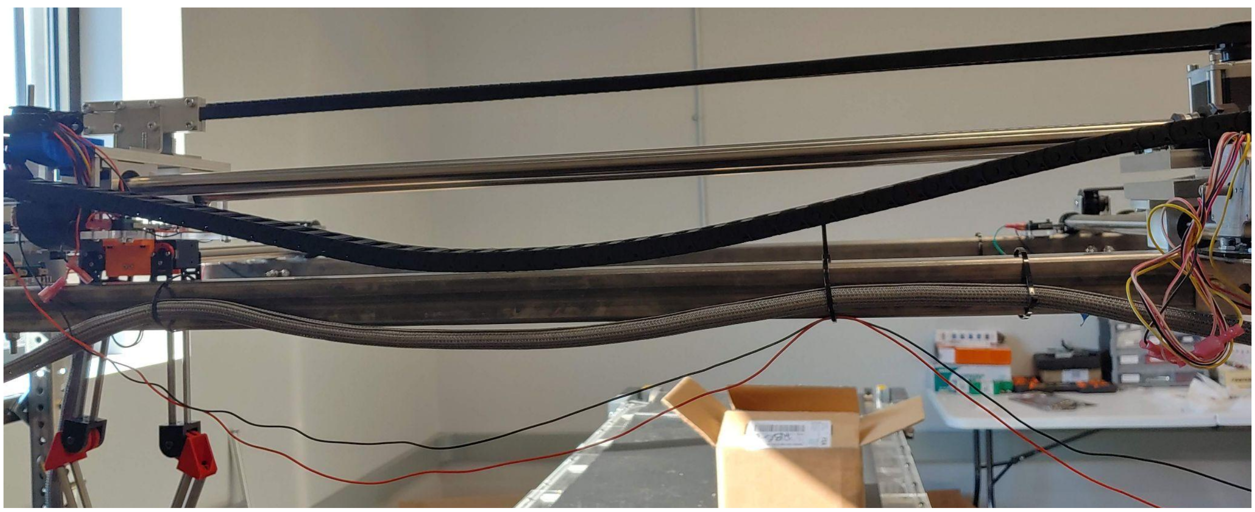


Figure 5.1: Limit Switch Wires Hanging off Robot Frame

Figure 5.2: Wire Harness Across the Robot Frame

While completing initial testing of the system, it was found that if the robot moves too far in the positive x-direction the black wire harness that houses the wires for the gripper gets caught outside the robot frame. This issue is shown in Figure 5.3 and makes it difficult for the end effector to traverse back in the negative x-direction. To fix this, we designed a new mount that securely attaches the end of the wire harness to the end effector and keeps the wire harness from moving side-to-side as the end effector moves down the linear rail (see Figure 5.4).

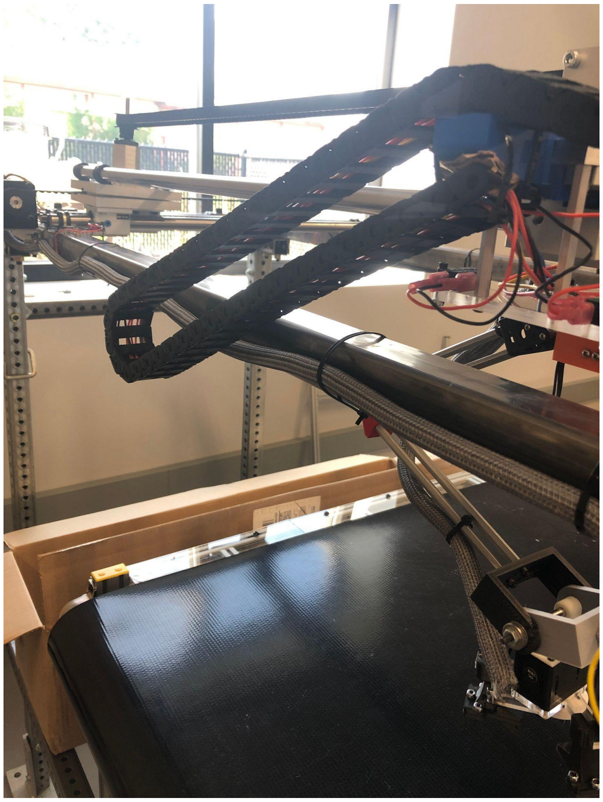


Figure 5.3: Wire Harness Stuck Outside the Robot Frame

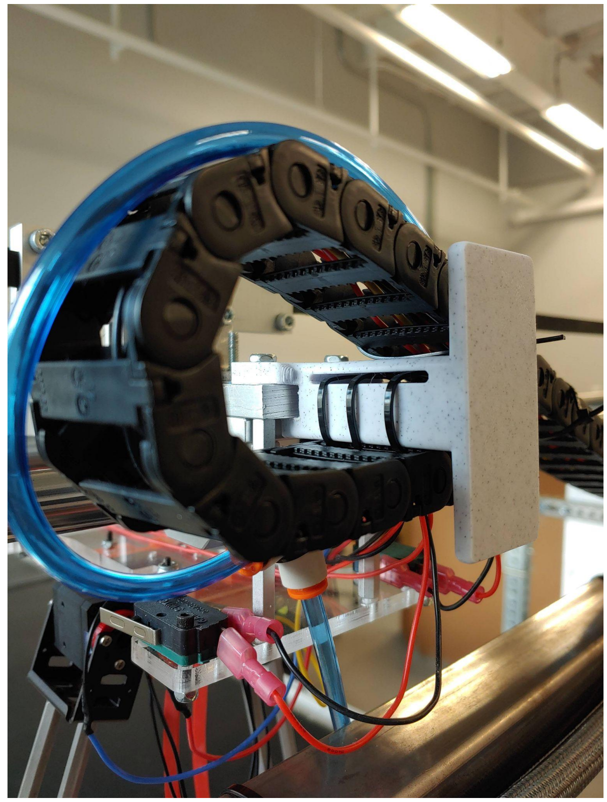


Figure 5.4: Updated Wire Harness Bracket

Linear Rails

When conducting our initial overview of the robot we learned that the linear rail had a couple different issues that needed to be addressed. The two main issues were making the carts that travel on the linear rails parallel and greasing the rails so the carts could move across them easier. First, we looked into how the carts work and what could be causing them to be misaligned. We learned that there is not anything connecting the upper cart to the bottom cart, instead they stay aligned because of individual countersinks on the upper cart that lineup with the screws that connect the belt to the lower cart. If these screws are not properly lined up with their respective countersink the carts are misaligned, as can be seen in Figure 5.5. To fix this we lifted up the upper cart and shifted it to realign the screws. This proved to be difficult because the upper cart was being kept in tension by the belt that moves the end effector in the x-axis. However, we were still able to align the screws and their countersinks to make the upper and lower carts parallel again (see Figure 5.6).

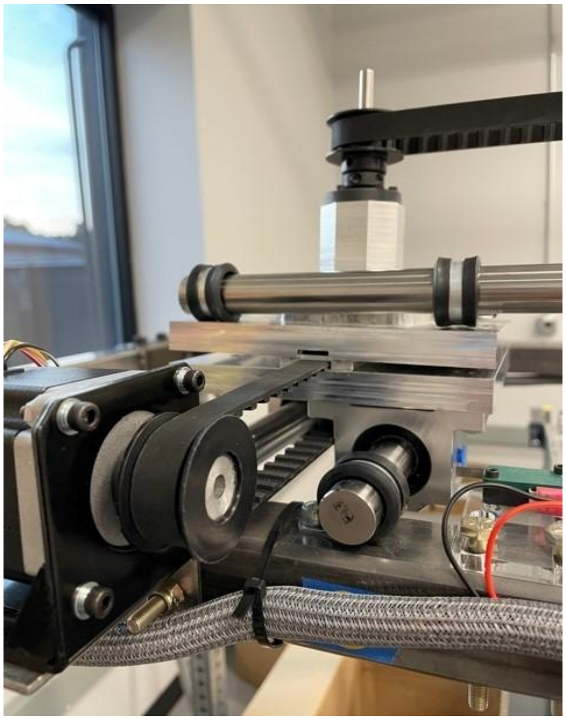


Figure 5.5: Misaligned Carts

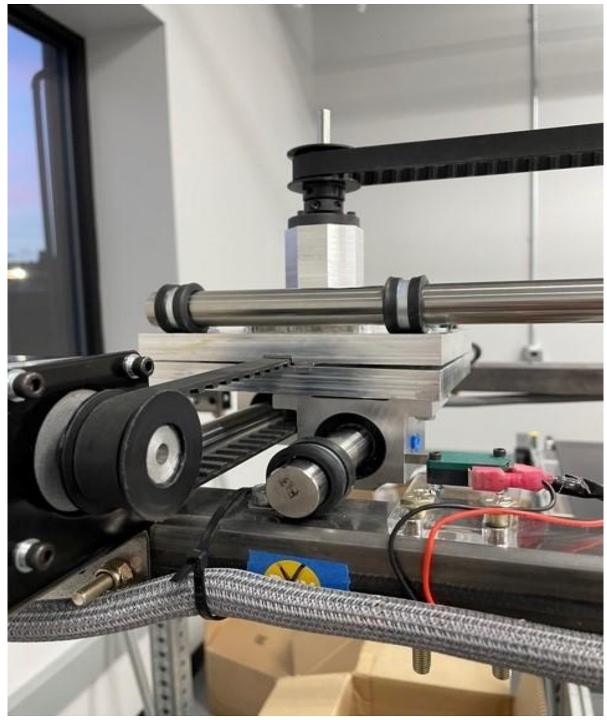


Figure 5.6: Aligned Carts

Realigning these carts helped the end effector to move smoother across the linear rails on the x-axis because they were previously not parallel because the carts were mounted at an angle. While this helped, we also decided that cleaning and greasing the rails would help because the rails were dusty. We used Krytox by Chemours GPL 204 Grease as recommended to grease the rails which seemed to help initially; however, the grease needs to be applied every two months to ensure the end effector can move easily across the linear rails.

5.1.2 Gripper

Selecting the Type of Gripper Mechanism

After researching the different types of grippers, we decided that designing a vacuum gripper would give us the results we wanted. A vacuum system is most suited for picking up flat materials compared to the other mechanisms we looked at. Even though cardboard is a porous material, with further research we discovered that vacuum systems can still pick the material up if the vacuum is running constantly. This differs from how vacuums work with non-porous materials, but ultimately is not a concern for our robot. We decided to move away from a mechanical gripper design because the original design did not prove strong or reliable enough to be able to pick up flat pieces of cardboard, much less any cardboard at all.

Initial Designs

Preliminary sketches were drawn up to brainstorm our potential design. Knowing that our main gripper would be a vacuum mechanism, we thought of different types of layouts that might suit our purposes best. We knew that placement of the vacuum cups would be important in the way we would be able to pick up cardboard. We thought of keeping the vacuum cups all in a line, as seen in Figure 5.7.

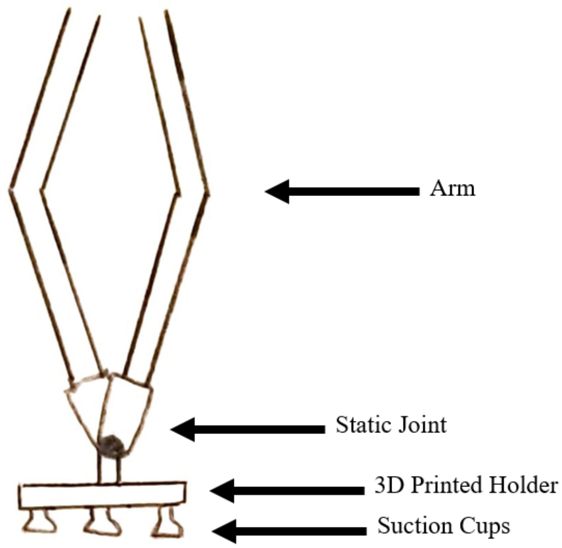


Figure 5.7: Basic Vacuum System Sketch

We also brainstormed vacuum gripper designs that included a mechanical mechanism. The design came from the potential need that we thought the mechanical mechanism might solve. The mechanical mechanism would be used to pierce the cardboard if it happened to have an irregular shape and the vacuum cups were unable to form a seal with any part of its surface. The different designs we came up with can be seen in Figure 5.8 and 5.9.

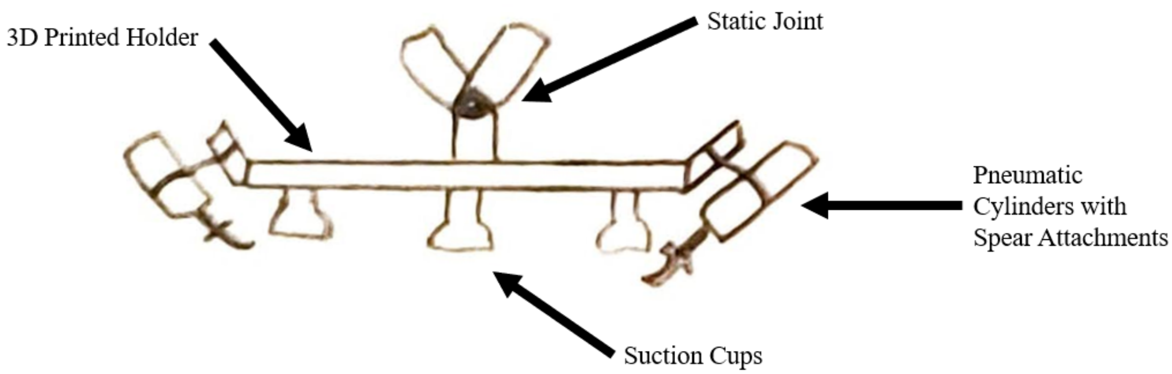


Figure 5.8. Vacuum System with Pneumatic Cylinder Spears

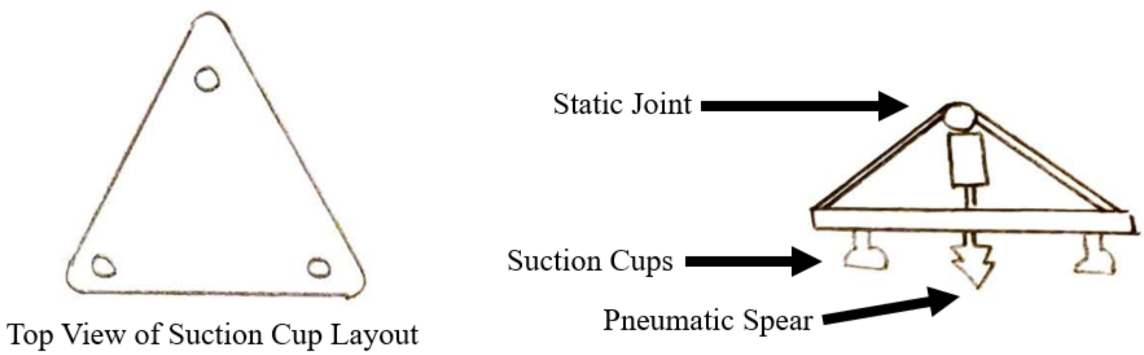


Figure 5.9. Triangular Vacuum System with Pneumatic Cylinder Spear

We thought of adding a spear mechanism to pierce the cardboard while the vacuum cups were trying to engage with the cardboard. The hope was that this would allow for more reliable attachment to the cardboard, but there were also many cons to this idea. We feared that the spearing mechanism would be too heavy for the arm to produce enough torque to be able to hold the gripper and the piece it would like to pick up. We also thought that it would prove difficult to let go of a piece after it has been picked up as it might get stuck on the spearing mechanism. There was also the possibility that the spear might pierce the conveyor belt if the cardboard piece was not big enough.

These brainstorming sessions led us to first test the vacuum system by itself and then decide if an added mechanical system was necessary. We decided to move onto the initial testing stage of our design process with two designs: three vacuum cups in a straight line and three vacuum cups in a triangle.

Initial Testing

We laser cut two different baseplates for the initial designs and assembled the vacuum system to them. An example of the setup we were using can be seen in Figure 5.10. Once we had assembled the different orientations, we had several different parameters that we were testing to determine which configuration to use:

1. Ability to pick up large pieces of cardboard.
2. Ability to pick up wet cardboard.
3. Ability to pick up boxes.
4. Ability to pick up flat cardboard.
5. Ability to pick up small pieces of cardboard.

After conducting these tests we found the following results:

- Both configurations struggled with picking up small pieces of cardboard without also grabbing the conveyor belt.
- The straight line configuration struggled when picking up boxes if the suction cups landed on the seam of the box because they could not form a tight seal.

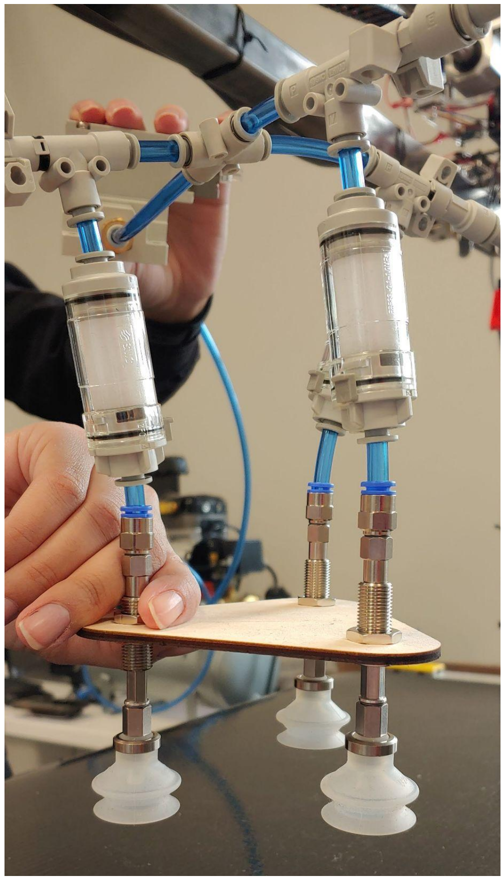


Figure 5.10: Initial Triangle Configuration for Testing

Ultimately we decided that the triangle configuration would be the best design to move forward with. However, we needed to make a few changes to the design so it could successfully pass all of our initial tests. We changed the heights of the three suction cups to have one cup lower than the other two. This cup would act as the primary gripper and the other two would act as supports when larger items were being picked up. This update to the design was made by changing where the suction cups screwed into the mounting plate (Figure 5.11).

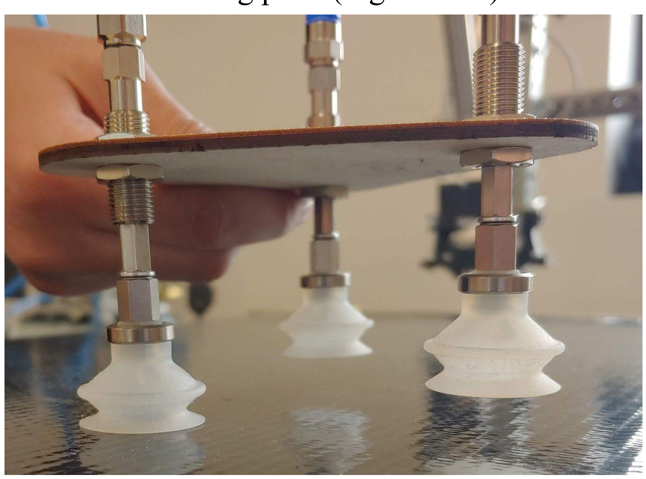


Figure 5.11. Updated Triangle Design with Primary Suction Cup

Re-Design

Once we had figured out the layout of the suction cups we worked on a design for the mounting plate. The mounting plate is the part of the gripper that connects the arm to the suction cup plate. Unlike the previous year's design, we separated the two parts of the scissor arm and had them attach to either side of the mounting plate as can be seen in Figure 5.12. The different holes in the mounting plate are for connecting the vacuum components. The assembled design can be seen in Figure 5.13; after assembling the gripper, we learned that it was too tall to fit between the end of the arm and the conveyor belt.

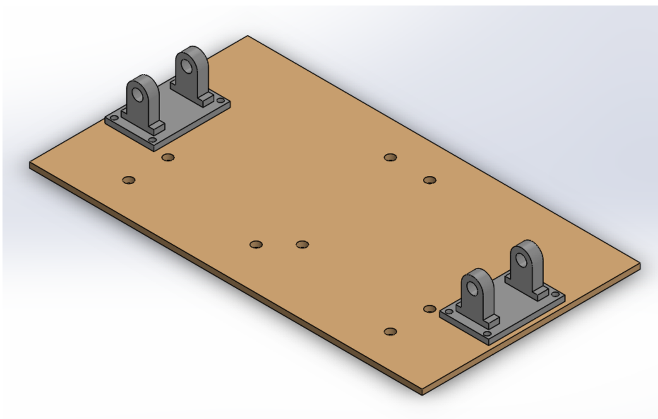


Figure 5.12: Mounting Plate Design

Figure 5.13: Assembled Design

To fix this issue, we came up with a new design that placed the filters above the mounting plate and thus shifted everything else up. This new design can be seen in Figure 5.14. The tolerances between the filters and the solenoid are very tight, but they allow for the design to connect to the arms and not interfere with the conveyor belt.

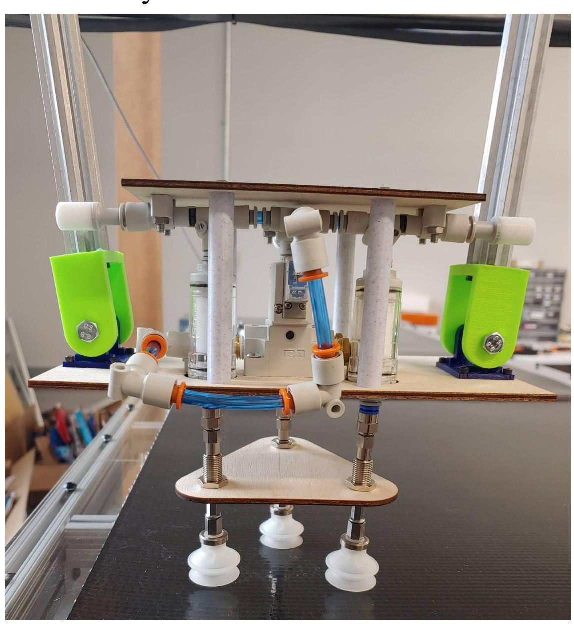


Figure 5.14: Updated Gripper Design to Fit Space Constraints

Through testing the robot with this updated configuration we learned that the servos used to lift the gripper are not strong enough to lift all the components needed for the vacuum system and an item that the gripper may pick up. To correct this, we moved the solenoid off the gripper to on top of the electrical box (Figure 5.15) and added a hook to the gripper to help the solenoids lift it (Figure 5.16).

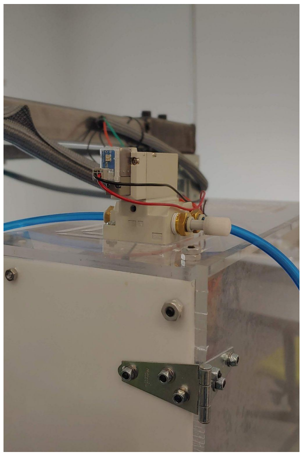


Figure 5.15: Updated Solenoid Position

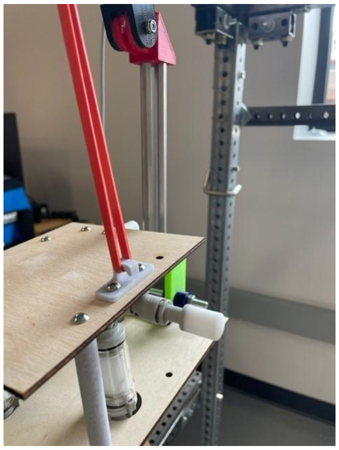


Figure 5.16: Added Hook for Weight Problems

Final Vacuum Design

After correcting all of these issues, we were happy with our final design and could start wiring the solenoid. The final gripper design can be seen in Figure 5.17 with the solenoid removed and the hook attached to the back of the upper plate.

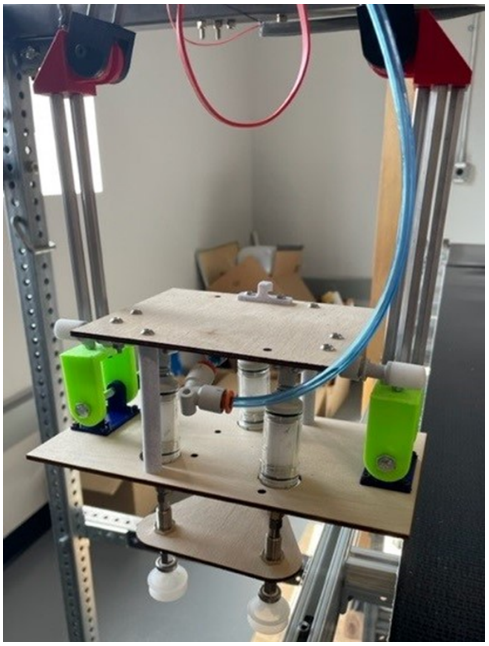


Figure 5.17: Final Gripper Design

We needed to buy a relay for the solenoid to convert our 24Vdc power source into 3.3Vdc. We ordered a TWTADE SSR-40 DD 40A DC Solid State Relay to help with this and wired the relay and solenoid into our electronics using the schematic in Figure 5.18. After testing the system to ensure everything was wired correctly we added the relay into our electrical box (see Figure 5.19).

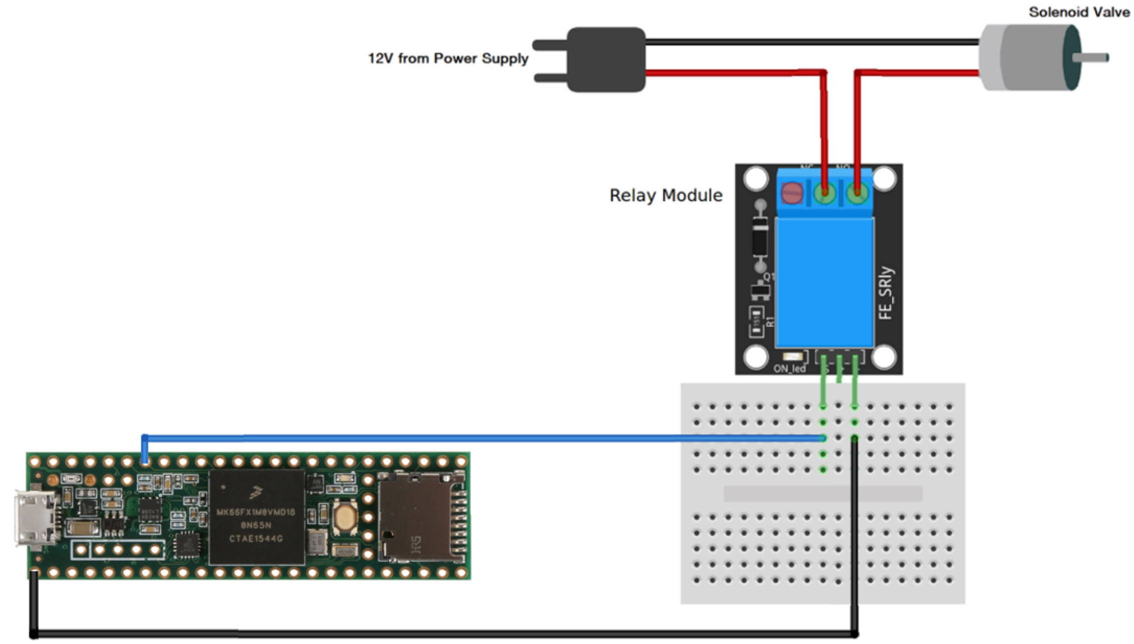


Figure 5.18: Wiring Schematic for the Solenoid

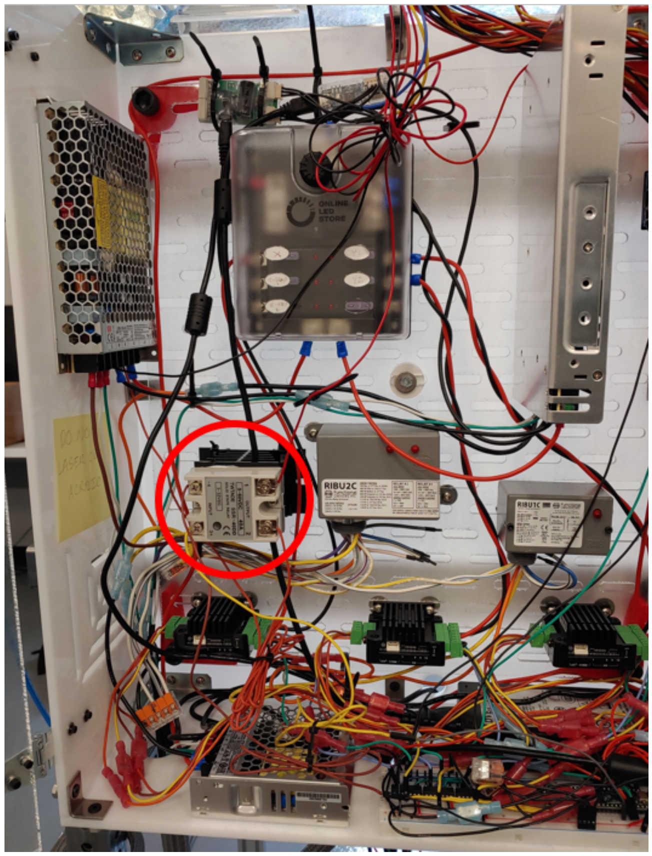


Figure 5.19: Relay within the Electrical Box

5.2 Software

5.2.1 Robot Control

For our final robotic system, we implemented 3 main ROS nodes to control all necessary functions. As seen in Figure 5.20, these 3 ROS nodes are the camera node, master node, and arm node. The camera node is responsible for object detection, object sorting, and human robot interaction (HCI). Each of these three aspects are covered in greater detail in the corresponding following sections.

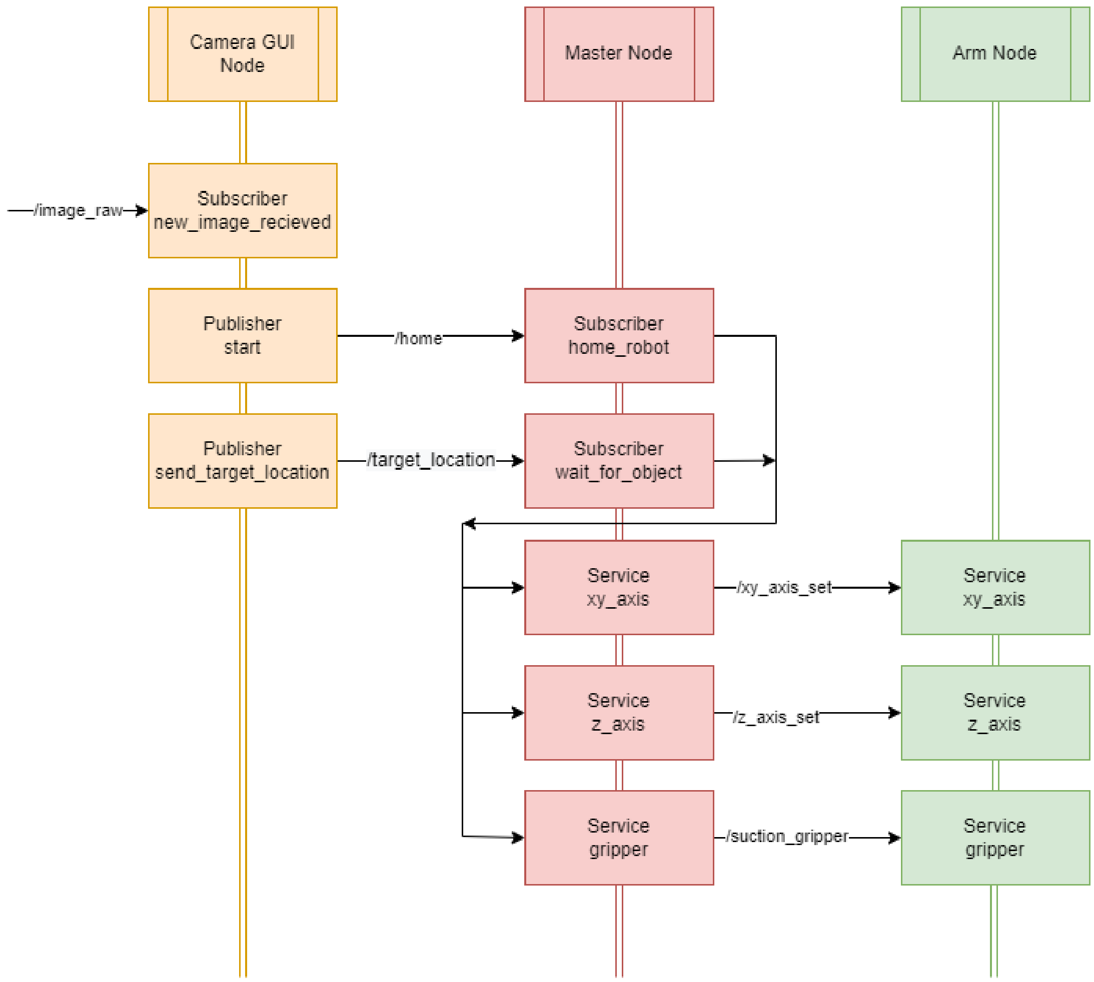


Figure 5.20: ROS Node Architecture

The arm node controls the mechanical functions of the robot through both the Dynamixel and Teensey. On setup, the node first ensures that both the Dynamixel and Teensey are running properly before opening up to commands. Specifically, the Dynamixel handles the 2 Dynamixel motors used to raise and lower the arm. The Teensey object in the node relays all commands to the C code running on the Teensey. The x-y limit switches, x-y stepper motors, and suction solenoid are all handled by the state machine on the Teensey as shown in Figure 5.21. The state machine begins by waiting for a homing command before accepting any other instructions. During the homing stage, the stepper positions are zeroed at the limit switches in a given corner. Once homed, the state machine waits idly until an instruction is given to either move the arm or engage or disengage the suction to the gripper. If any of the limit switches are triggered while moving, the current move command is terminated and the stepper motors are re-homed for the corresponding axis.

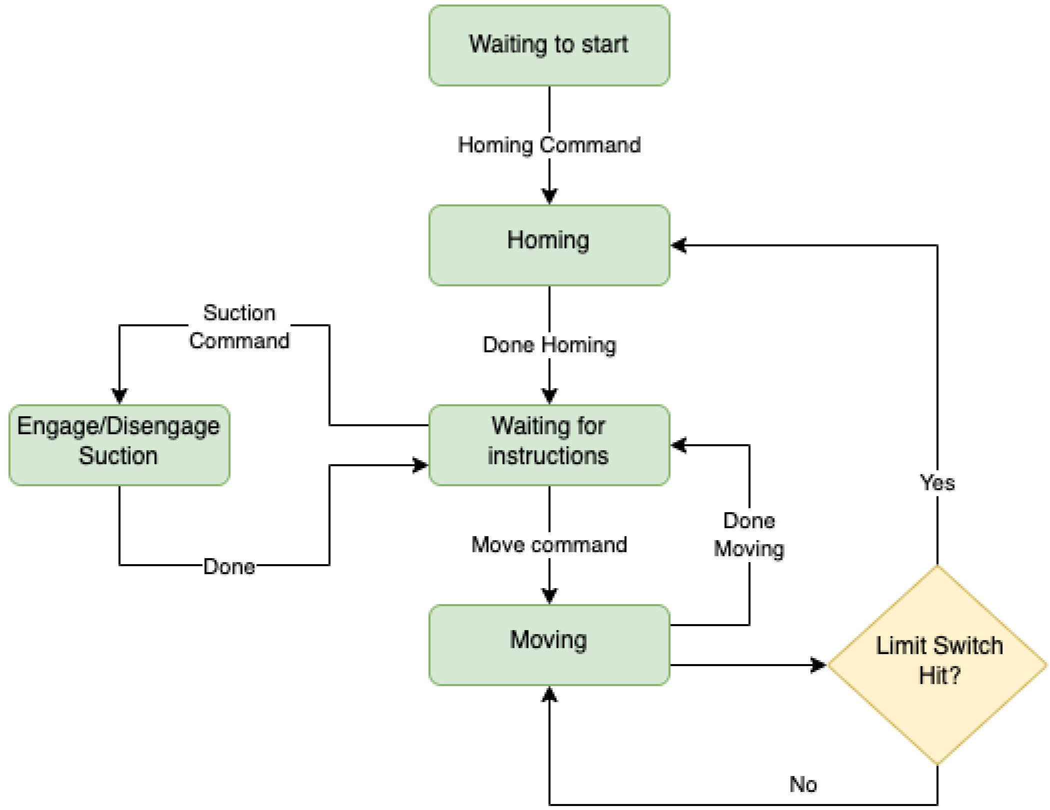


Figure 5.21: Teensey State Machine Logic

Finally, the master node facilitates communication between the GUI node and the arm node as depicted in Figure 5.20. When the Start button is pressed on the GUI, the camera node informs the master node to home the robot, telling the Dynamixel to raise the arm and the Tennsey to zero the stepper motors. Once the object detection and user interaction begins, whenever a target piece of cardboard is located, the camera node posts the current location and the master node begins a timed sequence. The master node will begin by preparing for the cardboard by moving the arm to the corresponding y position. Then, the node will wait until it predicts the cardboard is beneath the arm based on the initial cardboard position and the belt speed. The master node then picks up the cardboard by sending commands to engage the suction, drop the arm, and then raise the arm. Once picked up, the maser node will tell the arm to go to the dropoff position and disengage the suction before waiting for a new target cardboard.

5.2.2 Object Detection

Dataset Creation

To supplement the ZeroWasteAug dataset and fine-tune our object detection model, we created the Sagamore Lab dataset. With the goal of maximizing shared features that could generalize to both datasets, the Sagamore dataset has multiple factors consistent with ZeroWaste. Our dataset contains the same recycling classes as ZeroWaste: cardboard, soft plastic, rigid plastic and metal. Paper is classified as part of the background class just as it is in ZeroWaste. We also maintained roughly the same class distribution as ZeroWasteAug. The class distribution of both datasets are shown in Figure 5.22. While the Sagamore dataset has a smaller total number of annotations compared to ZeroWasteAug, the proportions of cardboard to other classes of recycling are approximately the same.

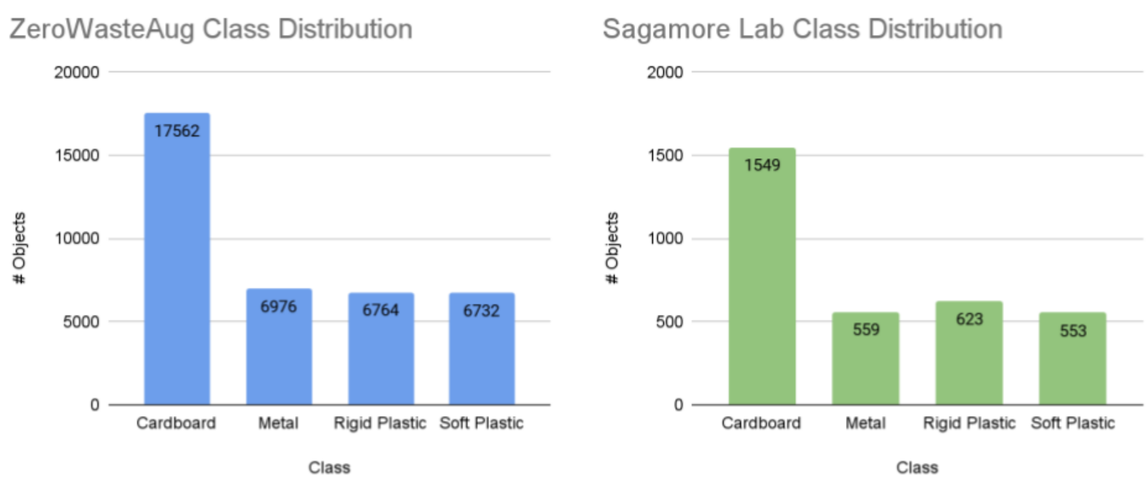


Figure 5.22: Class distribution of ZeroWasteAug (left) and Sagamore Lab dataset (right)

The images in the Sagamore dataset were collected with video footage recorded on an Intel RealSense Depth Camera D435 mounted above the conveyor belt in our lab. One frame for every second of the video was saved for annotation in the dataset. To reduce sampling bias, recycled materials were collected from seven different households. With the conveyor belt running at 22 centimeters per second, we placed pieces of recycling on the belt at varying positions and orientations. Sample images from the dataset are shown in Figure 5.23. To ensure separation between test and training data, the test dataset was created using a different collection of recycled materials than the training and validation sets.

Figure 5.23: Sample annotated images from Sagamore Lab dataset

Table 5.1: Comparison of Sagamore Lab dataset to ZeroWaste-f and ZeroWasteAug

	Sagamore Lab	ZeroWaste-f	ZeroWasteAug
# images	750	4,503	4,503
# annotations	3,284	27,468	38,034
Clutter (annotations/image)	4.4	6.1	8.4

Due to limited time and annotators, we produced a small dataset of 750 images with bounding box annotations. The Sagamore dataset has an average of 4.4 annotations per image which is 28% less clutter than what occurs in a real MRF as seen by ZeroWaste-f which has 6.1 annotations per image on average. Table 5.1shows the clutter level in the Sagamore dataset as compared to ZeroWaste-f and ZeroWasteAug.

Model Training & Selection

For our two chosen models, Faster R-CNN and YOLO v5s, we trained 3 model variants for each from 3 different training sets. The first configuration we tried was using ZeroWaste-Aug as the training set with COCO and default initial weights for Faster R-CNN and YOLO respectively. From this initial test we found promising results for YOLO with a significantly higher AP of 63.5% over Faster R-CNN's 41.9%. Once we finished annotating our Sagamore Lab dataset, we trained two more model variants for each using the unaugmented and augmented datasets. For both these variants we set the initial weights to the trained ZeroWaste-Aug model. Interestingly, we found that both Faster R-CNN and YOLO performed better on the unaugmented data, likely due to both models underfitting to the training set when using the augmented data. Overall, we found the YOLO v5s model trained on unaugmented Sagamore Lab data with ZeroWaste-Aug initial weights yielded the highest AP score of 68.8% at an IoU of 0.5:0.95.

We also tested the speed of the best performing Faster R-CNN and YOLO v5s on the lab computer. Due to the lab computer lacking a dedicated GPU, we had to run both models on the CPU. As expected, the YOLO model ran significantly faster than the Faster R-CNN model. YOLO v5s, with a model size of 14 MB, ran at a speed of about 3 frames per second. Faster R-CNN, with a model size of around 800 MB, ran at about 1 frame every 5 seconds. As a result, with a slightly better accuracy and a significantly faster runtime, we chose to implement our best YOLO v5s model into our system.

Additionally, to expand upon ZeroWaste's results, we also trained our initial ZeroWaste-Aug and Sagamore Lab models on ZeroWaste's ZeroWasteAug and ZeroWaste-f test sets. While these results were not as promising, we did find that YOLO outperformed Faster R-CNN in each configuration. All trained model variants and tested configurations can be found in Table 5.2 and Appendix A. The confusion matrix for the test results of the best YOLOv5s model are also shown in Appendix A.

Table 5.2: Average Precision (AP) scores for object detection results on Sagamore Lab test dataset for Faster R-CNN and YOLOv5 training configurations

	Initial Weights	*Train/Val Data*	*All*	*Cardboard*	*Metal*	*Rigid Plastic*	*Soft Plastic*
Faster R-CNN	COCO	ZeroWaste-Aug	41.90	48.29	45.92	21.05	52.36
	ZeroWaste-Aug	Sagamore Lab	65.45	72.59	65.66	51.60	71.94
YOLOv5	None	ZeroWaste-Aug	63.50	73.30	66.60	43.90	70.40
	ZeroWaste-Aug	Sagamore Lab	68.80	73.40	71.30	56.90	73.50

5.2.3 Object Tracking

The following section describes how recycling objects are tracked through each new frame of the live camera feed as they move down the conveyor belt. The process is summarized in Figure 5.24.

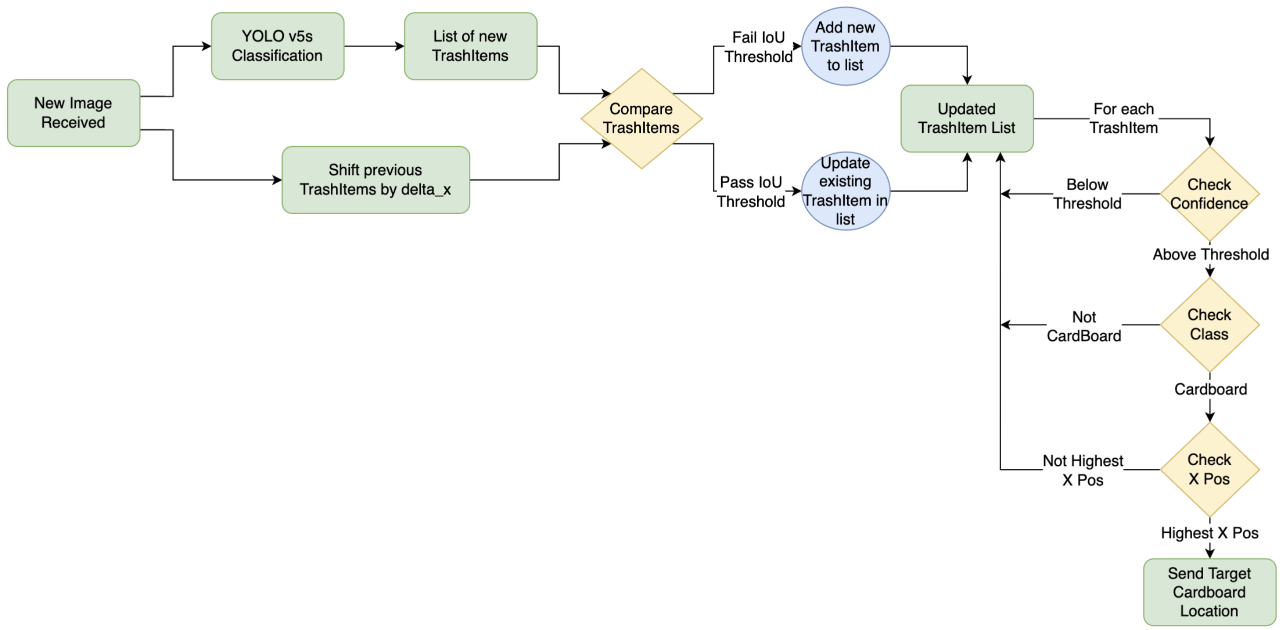


Figure 5.24: Object Tracking Procedure

When a new frame is received from the camera feed, it is run through the YOLO object detection model and a TrashItem object is created for each piece of recycling detected in the frame. A TrashItem stores information about the detected object including its center position, width, height, class, and confidence. All TrashItems detected in the previous frame are then shifted to the right by a distance delta_x which accounts for the distance an object would travel from one frame to the next at the given belt speed. This moves the objects from the previous frame into the same frame of reference as objects in the new frame. In order to properly update the x position of all TrashItems, it is necessary to account for the variable runtime of the YOLO model. Instead of adding a constant delta_x belt speed to each TrashItem, we calculate the number of frames that elapsed during the runtime of the model, and multiply that with a calculated distance per frame based on the belt speed. By re-computing delta_x for each new frame, we shift the x position of the TrashItem the correct distance, regardless of the speed of the model.

Once shifted, the previously detected TrashItems are compared with each newly detected TrashItem. If the intersection over union between two TrashItems is found to be above 0.5, the two objects are determined to be the same. The class, confidence, and position of the new TrashItem is updated into the previous TrashItem using Kalman filtering to compute an estimated confidence and y position. If a new TrashItem does not match any previously detected TrashItems, it is added to the TrashItems list.

The same TrashItem may be predicted as a different class from one frame to the next. Therefore, it is necessary to consider both the confidence and the number of frames an item is predicted in as a certain class to determine an accurate classification of the object. To track this information across all frames the object is detected in, each TrashItem maintains a set of Kalman parameters for each recycling class. Each set of Kalman parameters contains an array of confidences corresponding with each time the model predicted the object to be a given class, the current Kalman estimate of confidence for that class, and the error of the estimate. The class with the highest value of *confidence_estimate*num_frames* is selected as the classification of the TrashItem.

When an existing TrashItem in the list is updated with a new prediction from the current frame, the process is as follows. If the TrashItem has been manually created or updated by a user, only the x position of the item is updated. This allows us to accurately track the object's movement down the conveyor belt without overriding any human-made corrections. If the object has not been modified by a user, the Kalman estimate of confidence and y position are updated as well as the object's x position, y position, width and height. The Kalman parameters are updated using equations 1-4 below. For confidence, the set of Kalman parameters corresponding with the class predicted in the current frame are updated. After updating the confidence estimate, the class of the object is decided by which class has the highest $X\_EST\ *\ length(X)$.

$$\varepsilon_{mea} = error\ of\ measurements,\ \varepsilon_{est} = error\ of\ estimate,\ KG = Kalman\ Gain,$$

$$X = Measured\ values,\ X\_EST = Kalman\ estimate\ of\ X$$

(1) $$\varepsilon_{mea} = \sigma^2(X_{t=0:t})$$

(2) $$KG = \varepsilon_{est} / (\varepsilon_{est} + \varepsilon_{mea})$$

(3) $$X\_EST_t = X\_EST_{t-1} + KG * (X_t - X\_EST_{t-1})$$

(4) $$\varepsilon_{est} = (1 - KG) * \varepsilon_{est_{t-1}}$$

After all TrashItems have been updated or added to the list, a target TrashItem to be picked up is determined. For each TrashItem in the list, we check the confidence, class, and x position. To become the new target, the TrashItem must be classified as cardboard and have a confidence value above a threshold of 10. The confidence value is calculated by multiplying the confidence estimate by the number of detected frames for the item’s most confident class. For example, a piece of cardboard with confidence 0.9 will pass the confidence threshold after 12 detected frames. Out of all cardboard TrashItems above the confidence threshold, we select the item that is closest to the robot, having the highest x position, to be the target. The position of the new target item is then sent to the master node, so that the robot arm can move to the correct position to pick the item up.

5.2.4 Human Robot Interaction

With the goal of designing a user interface for human-aided recycling detection, we identified the following functional requirements:

1. The user can start and stop the recycling sorting system.
2. The user can see the bounding box, class, and confidence of all objects detected in the current frame.
3. The user can create new bounding box annotations for all recycling classes.
4. The user can edit the recycling class of any prediction currently displayed on the screen.
5. The user can delete any prediction currently displayed on the screen.
6. All corrections (create, edit and delete) can be made quickly.

These requirements ensure that the user can correct false positive and false negative predictions so that the robot can react according to the human-aid.

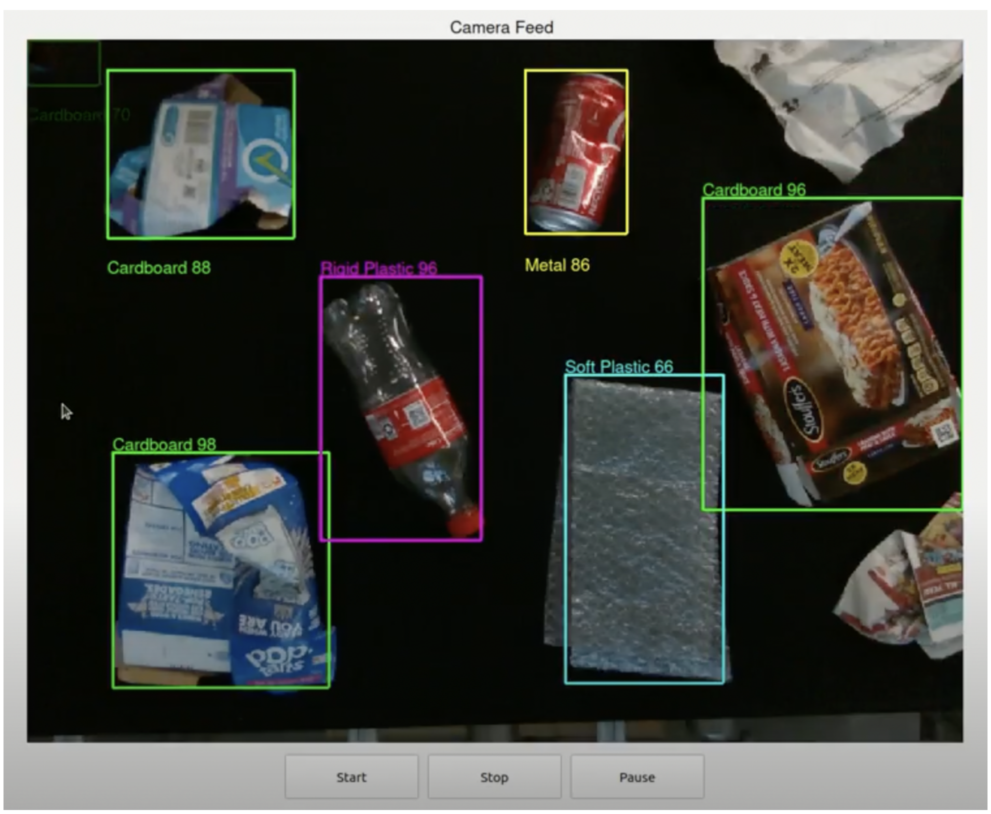


Figure 5.25: Screenshot of GUI with bounding box predictions of recycled materials

We implemented the user interface with Python's PyQt5 library. A sample screenshot of the GUI is shown in Figure 5.25. The GUI has a start button, which initiates the robot homing procedure and subscribes to the ROS camera node, and a stop button which unsubscribes from the camera node. Once the start button has been clicked, the live video feed is displayed on the screen and the object detection model is ready to make predictions for any objects that appear in the frame. As recycled materials move down the conveyor belt, the bounding box, class label, and confidence is displayed overtop the video for each object detected by the model above a certain confidence threshold. It is crucial to only display high confidence predictions so the user is not overwhelmed with predictions to correct. The low confidence predictions are not important to display, as they do not influence the robot's behavior.

To create new bounding box annotations, the user can click and drag to draw a box anywhere within the camera frame. Because the subject of the video is a moving conveyor belt, it is important to consider how challenging it would be for a user to draw a box around a moving object. Hence, the camera feed shown on the GUI is paused when a user initiates the drawing of a bounding box. With the video feed paused, the model continues running predictions and the sorting procedure in the background. After the user has completed the click and drag procedure to draw the box, a pop-up window appears, as seen in Figure 5.26, requiring the user to select a recycling class for the object. The user selects the class and saves the annotation. When the annotation is saved or canceled, the video feed becomes unpaused. The newly created TrashItem is then shifted over by a calculated distance that accounts for how far the object moved from the time the annotation was initiated to the time it was saved based on the speed of the conveyor belt.

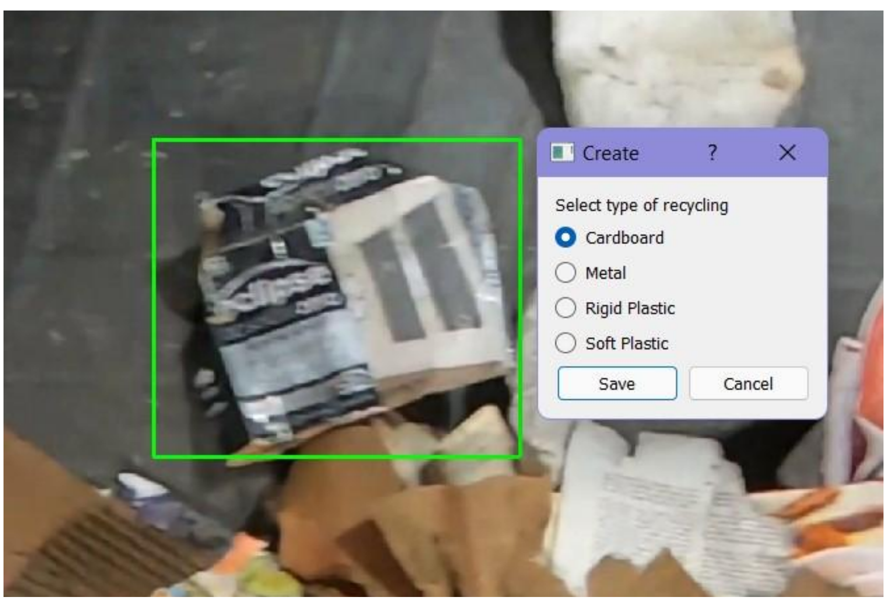


Figure 5.26: Pop-up for creating a labeled bounding box

To edit or delete any bounding box on the screen, the user must click anywhere inside the box to select it. If the user clicks within multiple overlapping bounding boxes, the smallest box will be selected. When a box is selected, a pop-up window is displayed, as seen in Figure 5.27. The user can select a different recycling class, boost the confidence to 100% for the current class label, or delete the bounding box entirely.

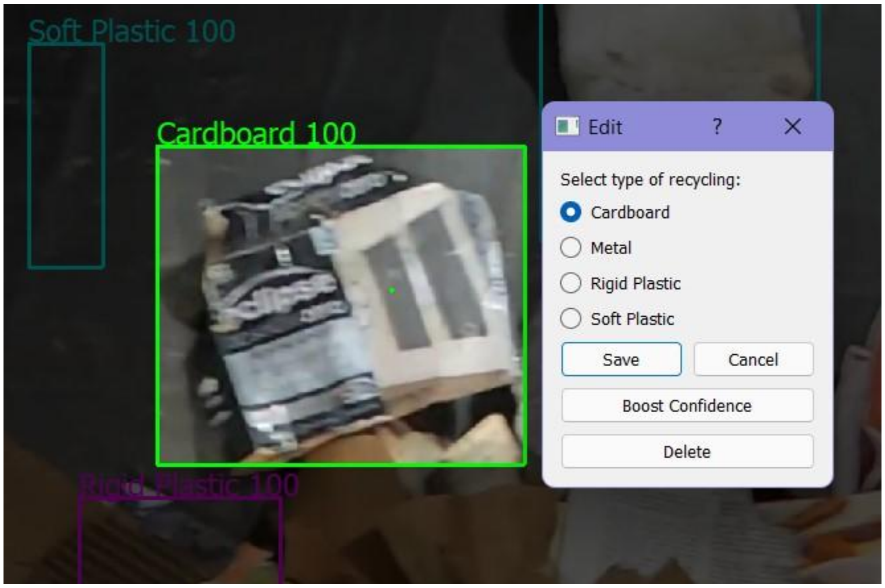


Figure 5.27: Pop-up for editing a bounding box label or confidence

After an object is created or modified by a user, the updated field stored in the TrashItem object is set to ‘True’. Any item flagged as being user updated will no longer be updated by new predictions from the model. This ensures the human-made corrections are preserved.

In addition to the basic functionalities described above, multiple improvements were made to the GUI to make it faster and easier to use. To make the user interface more clear, we use color to emphasize recycling type by displaying each bounding box and label in the color associated with the object’s class. Opacity is used to highlight confidence of the prediction. Boxes are drawn with higher opacity as the object is detected in an increasing number of frames. Boxes drawn at maximum opacity have met our confidence threshold of 10, as described in Section 5.2.3. As shown in Figure 5.27, when a user selects a box for editing, the background around the bounding box becomes darker to make it clear which box is selected. This is especially helpful when many boxes are present on the screen.

In order to make it easier to create and edit annotations, we also added a pause button. Although the video feed automatically pauses when a user begins drawing an annotation, it is challenging to click the correct position for the start of the bounding box while the object is moving. Consequently, a user may prefer to pause the feed before starting to draw an annotation. Similarly, when the user is trying to select a box for editing, it can be challenging to click inside a moving box especially if it is small. With a pause feature, the user can pause the video and then click inside the desired bounding box. We allow the user to pause the feed with the pause button or by hitting the spacebar. Once paused, annotations can be created or edited the same way as when the camera feed is running. To unpause the feed, the user must hit the spacebar or click the pause button again. The feed will also automatically be unpaused when a new annotation or edit is saved.

For faster user interactions, we optimized the placement of pop-up windows and added keyboard shortcuts for each action. The pop-up windows for editing and creating an annotation are always placed directly next to the selected bounding box on the side closest to where the user last clicked. This minimizes how far the user needs to move their mouse to interact with the pop-up. We implemented keyboard shortcuts for setting or changing the class of an object. When either the edit annotation pop-up or create annotation pop-up is open, the user can hit the keys 1, 2, 3, or 4 to automatically set the class and save it as cardboard, metal, rigid plastic, or soft plastic respectively. When a bounding box is selected for editing, the user can hit the delete key to delete the bounding box or hit 'B' to boost the confidence. Additionally, the user can pause or unpause the camera feed using the spacebar. While these shortcuts must be learned by the user, they greatly reduce the time needed to create and edit annotations, which is critical for making corrections fast enough to give the robot time to react.

5.2.4 Near IR

Adding an near infrared sensor to this project requires three steps. Firstly the assembly and calibration of a NIR camera. Secondly the creation of software to read and retrieve useful data from the NIR camera. Finally the full integration of the aforementioned software into the object tracking logic. This final step was not completed during this project and its discussion will be saved for the Discussion & Future Work section.

To begin we chose a NIR camera and appropriate accessories based on which would be most compatible with our project. We wanted an USB camera to match the camera and setup we had already worked with. In addition we wanted a camera that was capable of seeing only the NIR spectrum as having a camera that also reads the RGB would not only be redundant but would also make data difficult to read. This problem could be avoided by using a camera crafted to only sense NIR, but cameras like these are expensive and bulky [25]. Instead we opted to use a RGB + NIR camera, but this was initially an issue because we discovered that this camera did not sense all wavelengths equally (as seen in figure 5.28). The RGB spectrum of light is between 500 nm and 750 nm, while the NIR spectrum lies between 750 nm and 1200 nm.

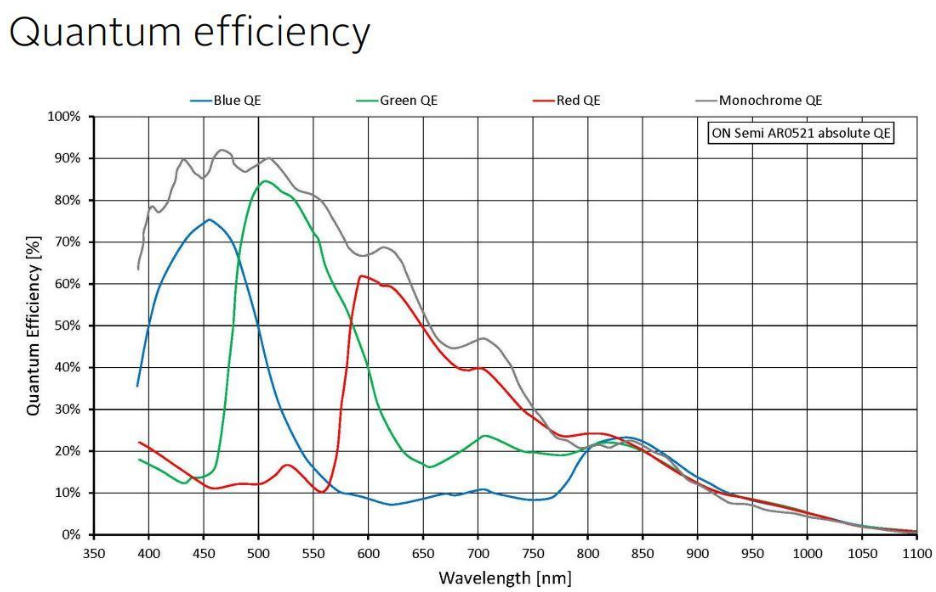


Figure 5.28: Quantum Efficiency for an RGB + NIR camera [23]

This figure shows that light of low wavelengths (sub 750) is much more likely to be picked up by cameras than light of large wavelengths (above 750) [23]. Since the camera reads much more RGB than NIR, the data would get easily drowned out by low wavelengths. This could be solved in two different ways: One, only fill the space where the camera is filming with NIR light; or Two, use a physical filter that reflects low wavelengths. Since we knew the camera would be set up in conjunction with a typical RGB camera, option one would be impossible. As such we decided to use a physical filter known as a Cold Mirror. This type of mirror reflects wavelengths lower than 750 nm which means the camera will only sense light in the NIR spectrum and provide us clean data [23].

Our total setup is as follows | Purchased from Edmund Optics [23]

- Camera - Allied Vision Alvium 1800 U-500m, 1/2.5" 5.0MP C-Mount, USB 3.1 Monochrome Camera
- Lens - 6mm UC Series Fixed Focal Length Lens
- Cable - USB 3.1 Gen 1 Micro-B to USB-A Locking Cable, 3m
- Lens Mount - 5mm Spacer to convert CS-Mount Cameras to C-Mount
- Cold Mirror - 0° AOI, 50.0mm Diameter, Cold Mirror
- Mirror Mount - M43 x 0.75 Mount for 50/50.8mm Diameter Filters
- Mirror Mount Adaptor - Filter Adapter M43 x 0.75 from 36mm Diameter

Mounts and adapters came at recommendation from Edmund Optics.

To set up and test the camera the Vimba SDK and Vimba Viewer application (both available through Edmund Optics) are required to calibrate and export settings for the Allied Vision Alvium camera. Once we set up the camera we used the VimbaPython library to grab images on demand for further analysis in python. Specifically we used blurring and contouring to isolate objects in an image before outputting their spectral analysis. Spectral analysis is simply just the breakdown of the different amounts of light reflected off of an object. Since our camera doesn't differentiate between different wavelengths of light, our spectral analysis consists of how much 750 - 1200 nm light each object reflects or brightness.

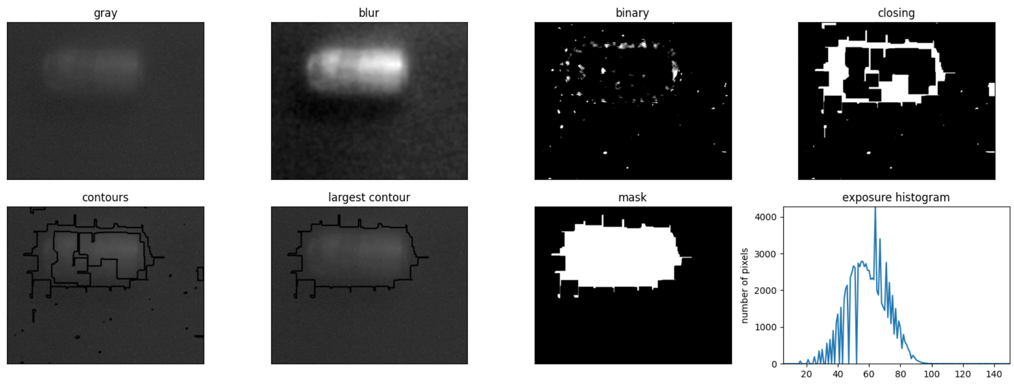


Figure 5.29: Sample output from our spectral analysis

Each image in Figure 5.29 describes a different step in the isolation process for objects we want to perform spectral analysis on. A mask is created for the object that isolates it from the background then a histogram of brightness values is produced. Since this histogram is produced by the NIR light reflected off the object it will provide information on the object material and not the color.

6. Overall System Integration and Evaluation

6.1 Experiment 1 - Clutter, Human Interaction, and System Accuracy

To evaluate our system as a whole, we performed 4 different types of trials as seen in Table 6.1, with 20 trials run for each type. Detailed trial-based results are shown in Appendix B. The goal for each trial was to correctly detect, pick up, and remove all cardboard from the conveyor belt amongst other objects. For the first two types of trials, each test consisted of 1 piece of cardboard and 2 pieces of non-cardboard cluttered closely together for a total of 3 objects. In trial 1, the system entirely relied on the YOLOv5s detection model for identifying cardboard. This was then compared to the addition of HRI in trial 2, where a user helped identify cardboard using the GUI. In trials 3 and 4, each test consisted of 1 piece of cardboard and 5 pieces of non-cardboard cluttered closely together for a total of 6 objects. Similar to the first 2 trials, trials 3 and 4 compared the results of not using HRI and using HRI respectively.

For the 4 trials in Table 6.1, we recorded 4 performance metrics. Detected total is the number of times the 1 piece of cardboard in each test was correctly identified. With the addition of HRI, we were able to increase the detected total from 75-80% to 100%. Given that the cardboard was correctly identified, hit total is the number of times the robot attempted to pick the cardboard. A reason the robot may not go for the cardboard is if the system also identified an object, like plastic, as cardboard and went for that instead. As such, using HRI to eliminate false positive cardboard readings, the hit total increased by 30-40% with the addition of HRI in both 3 object and 6 object trials. Picked up total is the number of times the robot successfully picked up the cardboard using the suction cups given that it the robot hit the piece of cardboard. Since this is mainly a mechanical performance metric, we saw a relatively equal performance across all trials at around 82-85%. Finally, the accuracy is the end-to-end performance of the entire system for each trial where the robot successfully identifies, hits, and picks up the cardboard. Overall, the addition of HRI increased the system accuracy from 55% to 70-80%.

Table 6.1: Percent detected, hit, and picked up for trials with and without human-aid at clutter levels of 3 objects and 6 objects

Trial Type	3 objects without HRI	3 objects with HRI	6 objects without HRI	6 objects with HRI
Detected Total	15 (75%)	20 (100%)	16 (80%)	20 (100%)
Hit Total	13 (87%)	19 (95%)	13 (81%)	17 (85%)
Picked Up Total	11 (85%)	16 (84%)	11 (85%)	14 (82%)
Accuracy	55%	80%	55%	70%

For the two trial types involving human-robot interaction, we recorded the time it takes to make each type of correction. For creating new bounding boxes, this is the time from when the user clicks to start drawing the box to the time when the annotation is saved. For update and delete, it is the time from when the user selects a bounding box to the time when the change is saved or the box is deleted. All trials were performed with the same user at the computer. The number of total interactions of each type and the average time of those interactions are shown in Table 6.2 below. We found that on average create, delete and update corrections take 0.170, 0.743, and 0.223 seconds respectively. The overall average time over all correction types was 0.241 seconds per interaction.

Table 6.2: Number of create, update, and delete user interactions and average time across 20 trials with 3 objects and 20 trials with 6 objects

Trial Type	3 objects		6 objects		Overall
Annot. Type	Count	Avg. Time (s)	Count	Avg. Time (s)	Avg. Time (s)
Create	4	0.242	9	0.139	0.170
Delete	1	0.083	6	0.853	0.743
Update	50	0.234	91	0.217	0.223
Total	55	0.232	106	0.246	0.241

6.2 Experiment 2 - Material Test

We decided to run additional testing on the vacuum gripper to evaluate how well the gripper could pick up other recyclable materials. The data for this test was gathered by running ten different trials for each type of material:

- Metal
- Soft Plastic
- Hard Plastic
- Paper
- Cardboard

For the test each material was sent down the conveyor belt and the gripper attempted to pick it up. Our tests consisted of us sending different types of recyclable materials one at a time on the conveyor belt and recording whether or not the gripper was able to pick up the material. We ran 10 trials with 5 different categories of materials. Materials were placed at random points along the belt throughout the trials. Our results can be seen in Table 6.3 and the trial-based results are shown in Appendix B.

Table 6.3. Results from Material Tests

Material	Picked up (%)
Metal	80%
Soft Plastic	100%
Hard Plastic	90%
Paper	70%
Cardboard	100%

7. Discussion & Future Work

7.1 Mechanical

7.1.1 Vacuum Gripper

From the results of our material test, we can conclude that the vacuum gripper is successful at picking up different types of materials as long as it can create a complete seal with at least one of the vacuum cups. Through conducting these tests, we learned that the vacuum has a 100% success rate with soft plastic which we were not expecting; this is because instead of creating a tight seal with the soft plastic the vacuum sucks up the soft plastic, especially plastic grocery bags. Since this happens, the gripper is not always able to drop the soft plastic because it gets stuck in the suction cups. It was unsurprising that paper had the lowest success rate for being picked up; this is mainly because most types of paper are too porous for a tight seal to be made. The rest of the results matched with our assumptions, we designed the gripper to be able to pick up cardboard, so we expected the pick up rate to be 100%. Also for metal and hard plastic we thought that the gripper would be able to pick up most of them, but when using cans and bottles the orientation of the object sometimes made it more difficult for the suction cups to form a tight seal which was expected.

This project could see further improvements with a gripper that is more versatile in being able to pick up all types of recyclable materials no matter their size, orientation, or material classification. One drawback to the vacuum is the fact that it has trouble picking up particularly porous materials, like paper. The addition of a mechanical mechanism that could work with the vacuum but also pierce the material to ensure accurate gripping could potentially help this. Initial designs of our gripper had looked to do this, but due to time and scope of our project was deemed unnecessary.

7.1.2 Overall Robot System

The overall robot would often move not as smoothly as we would have liked. Quick fixes like greasing the rails would sometimes help but were not a great long term solution. For future work on this project, a new rail system might help with smoother robot motion because currently the linear rails are not perfectly aligned which causes the system to be louder than is needed.. The linear rails were also speed limited which would at times cause the robot to miss a piece of cardboard that it needed to pick up because it wouldn’t have enough time to move to that detected piece. The improvement of the rail system might help this and/or the implementation of additional robot arms. Although it might be impractical for our prototype robot at the lab, this would definitely prove necessary at a real Material Recovery Facility. There would be too much waste to sort through for just one robot to work in this setting in general. Additional robots would improve accuracy and the time it takes to remove recyclables from the waste stream.

An additional thought that we had was the improvement of the robot arm. Instead of just three degrees of freedom, a robot with six degrees of freedom would be faster and more reliable. This would eliminate the need for the linear rail that we use on our system as the robots could work on either side of the conveyor belt to pick up and place recyclable materials from the stream. Using a six degree of freedom robot would also save space along the conveyor belt in which more robots could be placed for picking up cardboard or other recyclable materials.

7.2. Software

7.2.1 Robot Control

With the current system, we have a robust system for moving the robotic arm in all three axes, as well as controlling the suction cups to grip TrashItems. This system allows us to move to detected pieces of trash, pick it up, and move it off the conveyor belt. A useful feature that could be implemented by future teams is a way to control the speed of the conveyor belt. It would be useful to implement a system that can slow down the belt as more objects are detected to allow the arm adequate time to pick up all objects it identifies, and then be able to speed up as needed. We could also add a known distance checkerboard to the conveyor belt to be able to calculate the belt speed, which would adjust with the conveyor belt.

Additionally, when running the system, multiple files are needed to be run in unison, such as the camera node, GUI, and master controller. A launch file that combines all those files into one would make it easier to run the system.

7.2.2 Object Detection

Looking at the results found in Table 6.1, the combined success rate of successfully detecting and hitting cardboard for both trials without HRI was 65%. This means that 65% of tests the model both detected the cardboard and didn't detect non cardboard as false positives. While not an exact comparison, this success rate aligns closely with YOLOv5's test results of 68.8% overall accuracy found in Table 5.2. We believe the following issues could be addressed in the future to help increase the accuracy of the model.

One of the consistent issues we ran into while testing our YOLO object detection model was the lighting from different times of day. We gathered our initial dataset under a specific lighting environment, so our model was trained on a single brightness level. As a result, we had to manually adjust the exposure of the camera to roughly match the brightness of the training data during every test otherwise our model would yield poor classifications. Given more time, gathering training data under different lighting conditions or further looking into brightness data augmentation could help alleviate this problem. An algorithm to automatically set the camera exposure or a setting in the GUI could also be alternative solutions.

Another improvement that could be made to the object detection system is upgrading the lab computer to have a dedicated GPU. The current system uses the CPU for running the YOLOv5s model and struggles to maintain even 3 frames per second with an already small model. Upgrading the GPU would allow YOLOv5s to run up to 140 frames per second [14] or allow us to explore other deep learning models. For example, Faster-RCNN, and other similarly sized models, were too large to run on the current CPU, but would be viable with a dedicated GPU. Also, as mentioned in section 4.2.2, a RNN-CNN model could be a new type of model that may yield promising results with more computational power available.

7.2.3 Object Tracking

One of the major limiting factors to picking up TrashItems is the speed at which the system can navigate to the object, so it is important to add ways to speed up the system. One such way is to change how we decide whether or not to go for a trash item as it gets further down the conveyor. Currently our sorting system only goes for a piece of trash that has not passed a threshold of where the furthest away piece of trash can still be picked up by the gripper in time before it passes the arm. This can be improved by changing it to a variable threshold based on where the piece of trash is located on the conveyor belt. TrashItems that are closer to the home position of the gripper take less time to be grabbed, as there is less distance to travel to the object compared to TrashItems on the far side of the conveyor belt. This means TrashItems closer to the home position can have a lower threshold, and continue to be considered for being picked up further down the conveyor belt, as well as allow objects that are close together near the gripper home position to be picked up, when before the system wouldn't believe there was enough time.

7.2.4 Human Robot Interaction

To determine the effectiveness of our system's human-robot interaction component, we evaluated the speed of correcting object detections and the effect of HRI on the accuracy of the system. As shown in Table 6.2 from Experiment 1, on average it took 0.170, 0.743, and 0.223 seconds for create, delete, and update corrections respectively. Although create and update annotations were both fast, delete annotations were approximately 3x slower. As indicated by the user performing the corrections during our trials, delete annotations were slower due to the location of the delete key on the keyboard relative to other keys used for keyboard shortcuts in our user interface. A simple improvement would be to use the 'D' key as a shortcut for delete which is in close proximity to the number 1, 2, 3, and 4 keys used as shortcuts for modifying an object's class.

While the bounding box creations were the fastest of three correction types, in research, more optimized methods for interactive annotation tasks exist. For example, graph cut [31] approaches which predict object segmentations based on positive "object" and negative "background" clicks. While the approach is useful for segmenting complex objects in a minimal number of clicks, the simplicity of the click and drag approach is likely faster for our application. Because we only require a bounding box to determine the center of our object, the click and drag method of drawing a bounding box is both fast, at 0.170 seconds on average, and intuitive. Introducing an annotation method that involves prediction of what the user intends to draw would introduce another layer of complexity and potentially require additional corrections if the predicted annotation is incorrect.

In evaluating the impact of human-interaction on our sorting system, we found that human aid can substantially improve both detection accuracy and hit rate. As shown by the results of Experiment 1 described in Section 6.1, the addition of HRI increased the system accuracy by 20-30%. For both clutter levels, the detected rate was 100% with HRI meaning the user was able to create an annotation for the cardboard object if it was not already detected or correct the classification if it was misclassified. The hit rate, however, was 10% lower for the trials with 6 objects compared to trials with 3 objects. The cause of the lower hit rate is likely due to false positives where the robot identified a non-cardboard object as cardboard and attempted to pick that object up instead. This indicates that it can be challenging to correct or delete all false positive predictions in time, especially at a higher level of clutter. An added challenge occurs when an object gets detected right before moving out of the frame. In this scenario, the user has less time to react because the object moves out of the camera's view within a few frames of being detected. An improvement that could minimize this challenge would be to increase the confidence threshold for an object to be picked up. The threshold is a value that could be improved upon further tuning and raising it would give the user more time to correct a prediction before the object leaves the frame.

An additional feature that could be added to the GUI, is exporting frames in which the user makes a correction. These frames could be saved to later be reviewed by an annotator and reserved for future training of the object detection model. This would help increase the amount of data available in the recycling domain and target instances in which the model's predictions are incorrect.

7.2.4 Near IR

As seen in Figure 5.29 the jagged mask produced by our algorithm still needs to be tuned further. Through more testing a more accurate histogram of just the objects light scale could be produced. This would likely entail modifying the existing code to test many different parameter values until the optimal blur, binary and closing sizes were found.

Future work would also include integrating the current method of spectral analysis into the object detection logic. This would firstly require several experiments to discover the relationship between the histograms produced and the object's material. This could be done with machine learning or perhaps even simpler algorithms if the relationships found were obvious enough. By integrating NIR we would expect to see an increased performance in classifying objects.

8. Conclusion

The goal of our project was to build upon the previous year's MQP work of creating a conveyor belt and robotic arm to make a prototype for a waste sorting system. We designed and tested a new gripper which uses suction cups to grip different pieces of recycling such as cardboard, plastic and metal. We also developed our own recycling dataset and trained a deep learning model to identify pieces of recycling from an image. Using the information from the object detection model, we created a system to locate cardboard objects and pick them up. We also added human robot interaction where a user can create, delete, and edit bounding boxes and class labels of recycling objects for the arm to pick up. With further improvements, this system could eventually be used in a material recovery facility to help facilitate the separation of cardboard and other materials.

References

[1] G. Thung. Trashnet. https://github.com/garythung/trashnet, 2016.

[2] Pedro F Proença and Pedro Simoes. Taco: Trash annotations in context for litter detection. arXiv preprint arXiv:2003.06975, 2020.

[3] M. Koskinopoulou, F. Raptopoulos, G. Papadopoulos, N. Mavrakis and M. Maniadakis, Robotic Waste Sorting Technology: Toward a Vision-Based Categorization System for the Industrial Robotic Separation of Recyclable Waste. In IEEE Robotics & Automation Magazine, vol. 28, no. 2, pp. 50-60, June 2021, doi: 10.1109/MRA.2021.3066040.

[4] Bashkirova, D., Abdelfattah, M., Zhu, Z., Akl, J., Alladkani, F. M., Hu, P., Ablavsky, V., Calli, B., Bargal, S. A., Saenko, K.. ZeroWaste Dataset: Towards Automated Waste Recycling. In CoRR, vol. abs/2106.02740, arXiv preprint arXiv:2106.02740v3, 2022.

[5] Tsung-Yi Lin, Michael Maire, Serge Belongie, James Hays, Pietro Perona, Deva Ramanan, Piotr Dollar, and C Lawrence Zitnick. Microsoft coco: Common objects in context. In European conference on computer vision, pages 740–755. Springer, 2014.

[6] Tsung-Yi Lin, Priya Goyal, Ross Girshick, Kaiming He, and Piotr Dollar. Focal loss for dense object detection. In Proceedings of the IEEE international conference on computer vision, pages 2980–2988, 2017.

[7] Francisco Massa and Ross Girshick. maskrcnn-benchmark: Fast, modular reference implementation of Instance Segmentation and Object Detection algorithms in PyTorch. https: //github.com/facebookresearch/maskrcnnbenchmark, 2018.

[8] Yanghao Li, Yuntao Chen, Naiyan Wang, and Zhaoxiang Zhang. Scale-aware trident networks for object detection. In Proceedings of the IEEE/CVF International Conference on Computer Vision, pages 6054–6063, 2019.

[9] Liang-Chieh Chen, Yukun Zhu, George Papandreou, Florian Schroff, and Hartwig Adam. Encoder-decoder with atrous separable convolution for semantic image segmentation. In ECCV, 2018.

[10] Yuxin Wu, Alexander Kirillov, Francisco Massa, Wan-Yen Lo, and Ross Girshick. Detectron2. https://github. com/facebookresearch/detectron2, 2019.

[11] Wilson, H., & Daugherty, P.. How Humans and AI Are Working Together in 1,500 Companies. from https://hbr.org/2018/07/collaborative-intelligence-humans-and-ai-are-joining-forces, 2019.

[12] S. Ren, K. He, R. Girshick, και J. Sun, ‘Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks’, Proceedings of the 28th International Conference on Neural Information Processing Systems - Volume 1, Montreal, Canada, 2015, pages. 91–99.

[13] J. Redmon, A. Farhadi, “YOLOv3: An Incremental Improvement”, 2018.

[14] X. Zhu, S. Lyu, X. Wang, en Q. Zhao, “TPH-YOLOv5: Improved YOLOv5 Based on Transformer Prediction Head for Object Detection on Drone-captured Scenarios”. arXiv, 2021.

[15] C. Wang, J. Qin, C. Qu, X. Ran, C. Liu, en B. Chen, “A smart municipal waste management system based on deep-learning and Internet of Things”, Waste Management, vol 135, pages 20–29, 2021.

[16] Q. Zhang et al., “Recyclable waste image recognition based on deep learning”, Resources, Conservation and Recycling, vol 171, page 105636, 2021.

[17] K. He, X. Zhang, S. Ren and J. Sun, "Deep Residual Learning for Image Recognition," 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 770-778, doi: 10.1109/CVPR.2016.90.

[18] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens and Z. Wojna, "Rethinking the Inception Architecture for Computer Vision," 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 2818-2826, doi: 10.1109/CVPR.2016.308.

[19] A. Howard, N. Zhem, B. Chen, D. Kalenichenk, W. Wang, T. Weyand, M. Andreetto, H. Adam, ‘MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications’, 2017.

[20] Choudhari, Varad & Pedram, Rushal & Vartak, Sohan. (2021). Comparison between YOLO and SSD MobileNet for Object Detection in a Surveillance Drone.

[21] S. Paul, “Keras Documentation: Video Classification with a CNN-RNN architecture,” Keras, 05-Jun-2021.

[22] Ming Liang and Xiaolin Hu, "Recurrent convolutional neural network for object recognition," 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015, pp. 3367-3375, doi: 10.1109/CVPR.2015.7298958.

[23] “Edmund Optics,” *Edmund Optics Inc.*, 2022 [Online]. Available: https://www.edmundoptics.com/

[24] Fan, Liang, X., Huang, W., Jialong Zhang, V., Pang, Q., He, X., Li, L., & Zhang, C. "Real-time defects detection for apple sorting using NIR cameras with pruning-based YOLOV4 network." Computers and Electronics in Agriculture, pg 193, 2022.

[25] "Optical sorting equipment - tomra auto sort recycling equipment - NIR Technology," *Van Dyk Recycling Solutions*, 06-Jan-2022. [Online]. Available: https://vdrs.com/tomra-optical-sorting/. [Accessed: 06-Apr-2022].

[26] *Types of grippers used in manufacturing*. Universal Robots. (n.d.). Retrieved March 28, 2022, from https://www.universal-robots.com/blog/types-of-grippers-used-in-manufacturing/

[27] *Brown, E., Rodenberg, N., Amend, J., Mozeika, A., Steltz, E., Zakin, M. R., Lipson, H., & Jaeger, H. M. (2010). Universal robotic gripper based on the jamming of granular material. Proceedings of the National Academy of Sciences, 107(44), 18809–18814. https://doi.org/10.1073/pnas.1003250107*

[28] Gauvin, G. (2022, March 21). *Building a robot cell [part 3]: Getting started with vacuum grippers*. Manufacturing Automation, Simplified. Retrieved April 6, 2022, from https://vention.io/blogs/getting-started-with-vacuum-grippers-431

[29] *8 Types of End of Arm Tooling Devices for Automation Projects*. Keller Technology Corporation. (2018). https://www.kellertechnology.com/blog/8-types-of-end-of-arm-tooling-devices-for-automation-projects/

[30] Brown, E., Rodenberg, N., Amend, J., Steltz, E., Zakin, M. R., Lipson, H., & Jaeger, H. M. (n.d.). *Universal robotic gripper based on the jamming ... - pnas.org*. PNAS. Retrieved April 22, 2022, from https://www.pnas.org/doi/10.1073/pnas.1003250107

[31] N. Xu, B. L. Price, S. Cohen, J. Yang, και T. S. Huang, 'Deep Interactive Object Selection', CoRR, τ. abs/1603.04042, 2016.

[32] Yu Shan. "Sichuan: Overall Promotion of Domestic Waste Classification and Disposal," "四川:统筹推进生活垃圾分类和处置." 中华人民共和国中央人民政府, 15-Mar-2022. [Online]. Available: http://www.gov.cn/xinwen/2022-03/15/content_5679062.htm.

[33] "Policy documents," "政策文件," 中华人民共和国中央人民政府, 2022. [Online]. Available: https://www.mee.gov.cn/home/ztbd/2020/wfcsjssdgz/dcsj/wfcszcwj/.

[34] "2020 Sustainability Report," *Casella*, 2022. [Online]. Available: https://sustainability.casella.com/sites/default/files/2021-07/Casella-SustainabilityReport-2020.pdf.

[35] B. Gellerman, “How a New Policy in China Has Led to a Recycling Crisis in Mass,” *WBUR News*, 21-Mar-2019. [Online]. Available: https://www.wbur.org/news/2019/03/19/recycling-massachusetts-china-effect.

[36] B. Michael. “European Union Vs. the United States: Recycling Policies and Management,” *Youngstown State University / OhioLINK*, 2009. [Online]. Available: http://rave.ohiolink.edu/etdc/view?acc_num=ysu1243439695.

[37] D. A. Toto, “Continual evolution,” *Recycling Today*, 06-Oct-2021. [Online]. Available: https://www.recyclingtoday.com/article/agilyx-chemical-recycling-profile/. [Accessed: 15-Apr-2022].

[38] “History and Current State of Waste Management in Japan,” *Japan Environmental Sanitation Center*, 2014. [Online]. Available: https://www.env.go.jp/en/recycle/smcs/attach/hcswm.pdf.

[39] K. Matsuoka, “After Years of Dabbling, Japan Gets Serious About Plastics Recycling,” *Chemical & Engineering News*, 17-Oct-2021. [Online]. Available: https://cen.acs.org/environment/recycling/years-dabbling-Japan-serious-plastics/99/i38.

[40] “New Life to Plastic,” *Greenplast*, 2022. [Online]. Available: https://www.greenplast.org/en/.

[41] P. Dobrowolski, “Recycling in China: From zero to hero?,” *Waste Management World*, 20-Apr-2021. [Online]. Available: https://waste-management-world.com/recycling/recycling-in-china-from-zero-to-hero/.

[42] “Population,” *United Nations*. [Online]. Available: https://www.un.org/en/global-issues/population.

[43] R. Cairns, M. Maruyama, and J. Ogura, “This 'Zero Waste' Japanese Building is Made From 700 Donated Windows,” *CNN*, 25-Jan-2022. [Online]. Available: https://www.cnn.com/style/article/kamikatsu-zero-waste-center-japan-climate-hnk-spc-intl/index.html/.

[44] R. P. Lee, B. Meyer, Q. Huang, and R. Voss, “Sustainable waste management for zero waste cities in China: potential, challenges and opportunities,” *Clean Energy*, Volume 4, Issue 3, Sep-2020, Pages 169–201. [Online]. Available: https://doi.org/10.1093/ce/zkaa013

[45] T. Jennifer, “Waste Management in the Shadows in China.” *Wilson Center,* 4-Aug-2021. [Online]. Available: https://www.wilsoncenter.org/event/waste-management-shadows-china.

[46] W. Xing, “Sichuan-Tibet Expressway: An Important Link for Sino-Tibetan Exchanges,” *For the Belt and Road*, 13-Jul-2021. [Online]. Available: https://www.seetao.com/details/96018.html.

[47] “Waste and Recycling,” *Worcester Energy*, 2022. [Online]. Available: https://www.worcesterenergy.org/take-action/-waste-and-recycling#:~:text=Worcester%20currently%20recycles%20approximately%2032,up%20recycling%20materials%20for%20free.

[48] “Zero-sort Recycling,” *Casella*, 2022. [Online]. Available: https://www.casella.com/services/recycling/zero-sort-recycling.

[49] “Zero Waste Case Study: Seattle,” *Environmental Protection Agency*, 19-Jul-2021. [Online]. Available: https://www.epa.gov/transforming-waste-tool/zero-waste-case-study-seattle#:~:text=Seattle%20recycled%2056.9%25%20of%20its,the%201998%20low%20of%2026.8%25.

[50] “Kamikatsu,” *Kamikatsu Town*, 2022. [Online]. Available: http://www.kamikatsu.jp/.

[51] “Kamikatsu Zero Waste Center,” “上勝町ゼロ・ウェイストセンター,” *WHY*, 2022. [Online]. Available: https://why-kamikatsu.jp/.

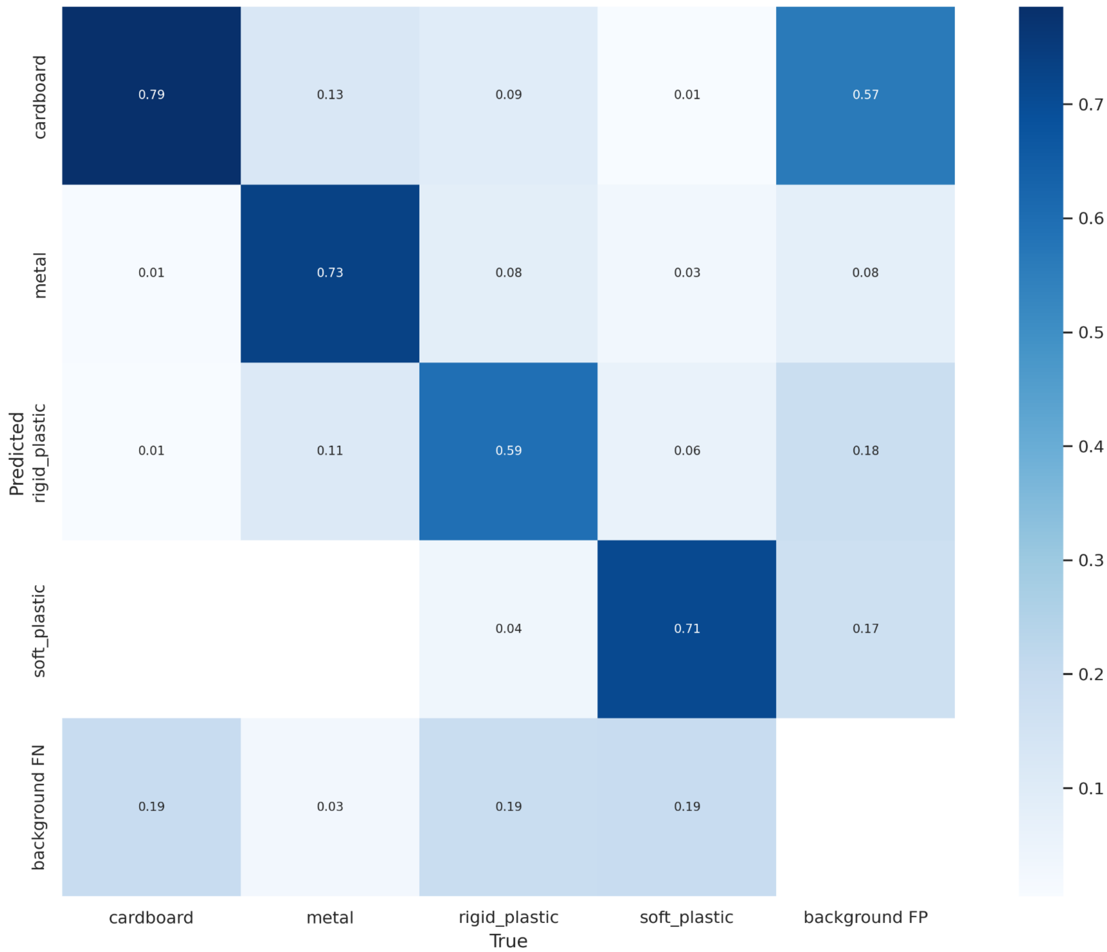
Appendix A - Object Detection Results

Table 5.2A: Average Precision (AP) scores for object detection results on different training configurations for Faster R-CNN and YOLOv5 models

Initial Weights	*Train/Val Data*	*Test Data*	*All*	*Cardboard*	*Metal*	*Rigid Plastic*	*Soft Plastic*
Faster R-CNN							
COCO	ZeroWaste-Aug	ZeroWaste-Aug Test	38.78	33.31	50.56	48.17	23.08
COCO	ZeroWaste-Aug	ZeroWaste-f Test	17.05	31.57	1.94	11.69	23.02
COCO	ZeroWaste-Aug	Sagamore Lab Test	41.90	48.29	45.92	21.05	52.36
ZeroWaste-Aug	Sagamore Lab	Sagamore Lab Test	**65.45**	**72.59**	**65.66**	**51.60**	**71.94**
ZeroWaste-Aug	Sagamore Lab-Aug	Sagamore Lab-Aug Test	60.19	67.10	60.60	46.98	66.08
ZeroWaste-Aug	Sagamore Lab	ZeroWaste-Aug Test	24.33	17.68	39.30	26.15	14.21
ZeroWaste-Aug	Sagamore Lab	ZeroWaste-f Test	11.35	19.83	1.50	9.47	14.57

YOLOv5							
None	ZeroWaste-Aug	ZeroWaste-Aug Test	55.10	39.50	74.20	67.40	39.20
None	ZeroWaste-Aug	ZeroWaste-f Test	26.80	32.00	1.96	38.70	34.40
None	ZeroWaste-Aug	Sagamore Lab Test	63.50	73.30	66.60	43.90	70.40
ZeroWaste-Aug	Sagamore Lab	Sagamore Lab Test	**68.80**	**73.40**	**71.30**	**56.90**	**73.50**
ZeroWaste-Aug	Sagamore Lab-Aug	Sagamore Lab-Aug Test	57.60	65.40	67.30	35.30	62.40
ZeroWaste-Aug	Sagamore Lab	ZeroWaste-Aug Test	35.50	23.10	53.90	42.90	22.30
ZeroWaste-Aug	Sagamore Lab	ZeroWaste-f Test	21.30	15.60	0.00	47.50	22.10

Figure 5.2A: Confusion matrix for best performing YOLOv5s model:



Appendix B - Experimental Results

Table 6.1B: Trial-based results of Experiment 1 described in Section 6.1

	Clutter	3 items per frame			6 items per frame		
	Trial #	Detected	Hit object	Picked up	Detected	Hit object	Picked up
No HRI	1	x	x		x	x	
	2	x	x	x	x	x	x
	3	x	x	x	x	x	x
	4	x	x	x	x	x	x
	5	x	x		x		
	6	x			x	x	x
	7				x	x	x
	8	x	x	x	x	x	x
	9	x	x	x	x	x	x
	10						
	11	x	x	x	x		
	12	x	x	x	x	x	x
	13				x	x	
	14	x	x	x	x	x	x
	15	x					
	16	x	x	x	x		
	17	x	x	x	x	x	x
	18				x	x	x
	19						
	20	x	x	x			
Total		15	13	11	16	13	11
% of Possible		0.75	0.87	0.85	0.8	0.81	0.85
Overall Accuracy		0.55			0.55		
HRI	1	x	x		x	x	x
	2	x	x	x	x		
	3	x	x		x	x	x
	4	x	x	x	x	x	x
	5	x	x	x	x	x	x
	6	x			x	x	x

	7	x	x	x	x	x	x
	8	x	x	x	x	x	x
	9	x	x	x	x	x	x
	10	x	x	x	x	x	x
	11	x	x	x	x	x	
	12	x	x	x	x	x	x
	13	x	x	x	x	x	x
	14	x	x	x	x	x	x
	15	x	x		x	x	x
	16	x	x	x	x	x	
	17	x	x	x	x		
	18	x	x	x	x	x	
	19	x	x	x	x	x	x
	20	x	x	x	x		
Total		20	19	16	20	17	14
% of Possible		1.00	0.95	0.84	1.00	0.85	0.82
Overall Accuracy		0.80			0.70		

Table 6.3B: Trial-based results of Experiment 2 described in Section 6.2

Trial #	Metal	Soft Plastic	Hard Plastic	Paper	Cardboard
1	1	1	1	1	1
2	1	1	1	1	1
3	0	1	1	0	1
4	1	1	0	0	1
5	1	1	1	1	1
6	1	1	1	1	1
7	0	1	1	1	1
8	1	1	1	0	1
9	1	1	1	1	1
10	1	1	1	1	1
Pick %	80%	100%	90%	70%	100%