Garbage Classifier

—

Anirudh Vijayan

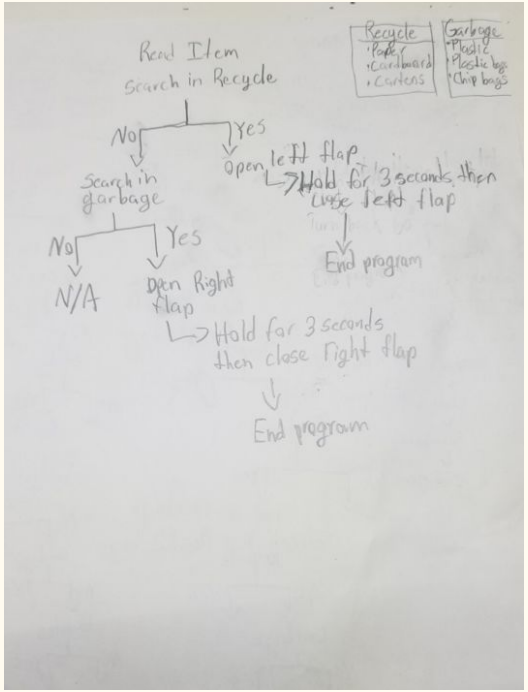
Question:

Can a garbage classifier positively affect the environment by ensuring that waste gets sorted properly?

Hypothesis

If I make a web-based garbage classifier, it will have a positive impact on the waste problems our planet is facing. Humans tend to put trash in the wrong bin, and all waste (recyclables and compost) ends up in landfills. This has a negative impact on the environment and makes it difficult to sort it out in the final stage. Making sure our waste is disposed of properly would decrease the waste in landfills.

Planning Sheet



This is the first planning sheet I have made, it is showing every step in detail of what I want my smart dustbin to do when different types of garbage are placed in it. The computer won't know what the item is, so I would put all the different types of trash in a random order but in its section and the code I will make, will scan different types of items in the order I have made it in. Then it will use object detection to find the item I have placed in the smart dustbin. Then the computer will see which section it is in (garbage or recycle) and will ask itself, is it in recycle, if the answer is yes then it will do as I have written. Open the left flap (because I am planning to put the recycling portion on the left side), hold for three seconds (so the recyclable has enough time to fall into the bin), tilt back to normal, and end the program. If I am able to do this method then I can try to move onto a more complex method which involves having three different sections (recycle, garbage, and bottles). For this, the flap method won't work and will need a turning method which involves much further code and angles (for example, 30 degrees right) to get to the different sections.

Background Research

Trash today

In Canada, there are too many people putting trash in the wrong bin, most simply out of laziness to put it in the right bin or just not knowing which bin it goes in. 84% of households put things in the wrong bin. This is costing recycling programs around the country millions of dollars. One in three pounds (0.45kg in 1.36kg) of trash put in the recycling bin actually shouldn’t be there. Cities in Canada with very dirty recycling (Edmonton and Toronto) can have contamination increases by 25 percent. It's very expensive to process contamination as recycling, it can even cost up to $4 million for a city to do this. All this has become a big issue because China, which is the biggest importer of recyclables, banned importing 24 different types of waste to prevent environmental disasters in the country. One of the waste items they banned from importing was paper and this became a problem for the rest of the world. "Something as simple as a piece of paper with a coffee stain on it, that piece of paper a year ago would have been recyclable, Today that's actually garbage". These are words from Jim McKay because of the China importation ban.

Background Research

Recycling stats

In Calgary, 80 percent of what's in the recycling bin is recyclable, the other 20 percent is contaminated. Everything that is recycled goes to a factory to get sorted out. It goes through many stages before finishing. 70 percent is sorted by machine and 30 percent is manually sorted out. To remove metal, there is a large magnet that picks all the metal up and removes it from the lines. It also sorted out all the other recyclables like glass, paper, and cardboard. Whatever the machines left are then manually sorted leaving only materials that aren't recyclables. There are some flaws though, anything like styrofoam, large scrap metals, or garbage can interfere with the machines and cause them to not work. In Toronto, 70 percent of items in residential recycling bins are recyclable, while 30 percent are contaminated. That is 10 percent more than Calgary.


80%
of the items you put in the blue cart
are sorted and recycled
20%
is contamination
(items that do not belong
in the blue cart)

Background Research

What can and can’t be recycled

Plastic usually cannot be recycled, but if it is stretchy like grocery bags or bubble wrap, then it can! Containers made of plastic are also recyclable, a case where plastic wouldn't be recycled is if it's crinkly like a chip bag. If containers are made of tin, then they can be recycled, as well as tin foil and pop cans. Glass can also be recycled, in the factory, it will be crushed and shipped to companies to make new products. What can’t be recycled though is household items such as small appliances, trays, furniture, or lightbulbs. Toys and sports equipment can not be recycled either, along with clothing or shoes. Styrofoam will damage the machines in the factories so they cannot be recycled. Lastly, no hazardous materials as it is expensive to decontaminate.

Background Research

Trash Segregation problems today

The main reason I am doing this project is because of trash segregation problems. This is a problem because dangerous stuff like needles, dead animals, and bear spray are put in the recycling and it costs Canada lots of money to decontaminate them and all that money is wasted because it all ends up in the landfill either way. It costs Canada millions of dollars to decontaminate and segregate trash. Calgary's contamination rate for residential recycling is 13% of all of Canada and the only cities with a higher contamination rate are Edmonton, Toronto, Halifax, and Fredericton. Toronto's contamination rate is the highest at 26% because trash segregation is the worst over there. This is all just because of people who put the wrong things in the recycling, the process of sorting trash and decontaminating items thrown in the wrong bin costs a lot more than we would expect. In Toronto itself, it costs about $600,000 - $1 million per year, that's a lot just to decontaminate trash. This is why segregating trash is very much needed and otherwise Canada is going to keep wasting money on just segregating trash.

Background Research

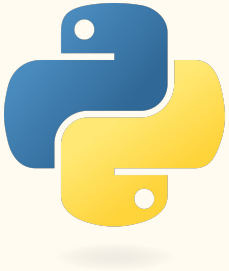
What is machine learning?

Machine learning is a strong sense of computer science that uses data algorithms to enable AI. Machine learning slowly and gradually learns from its mistakes and improves its accuracy. Based on input data, machine learning will give a prediction, and will produce an estimate of patterns in the data. In my case, it will use an error function, where it identifies a margin of error to decide how accurate the following piece of trash is to the picture stored in the image library. Machine learning is adapting to human behaviors and trying to imitate them. It uses various image sources to define which segregation it belongs to, kind of like how we work, we see the trash with our eyes, and using previous knowledge, put it in the correct segregation. That's how it learns over mistaken attempts, when it makes a mistake, it will acknowledge that and fix it, just like humans!

Background Research

What is python and streamlit?

Python is a form of coding language, it's very advanced and already has a lot of functions built into it. I will be using this to code to tell what I want my dustbin to do, and when to do it. I will also use classification where I will classify which segregation the piece of trash will end up in. Streamlit is a Python framework that delivers database apps with a few lines of code. It is what shows on the front end and is controlled with code in the backend. Streamlit creates data science apps mainly for machine learning. It’s a Python library that gets created fast without knowledge of web development.



Background Research

What is tensorflow?

Tensorflow is an open-source platform that software developers use heavily. It is mature in deep learning frameworks and can develop advanced models. It can be used to train models and datasets to make it easy to classify items. Tensorflow is based on machine learning and the accuracy increases after the model training. The framework will input the data as multi-dimensional arrays (arrays with more than one dimension), these are called tensors. It is used for many tasks, such as image recognition, language processing, handwriting recognition, and certain equations.

TensorFlow

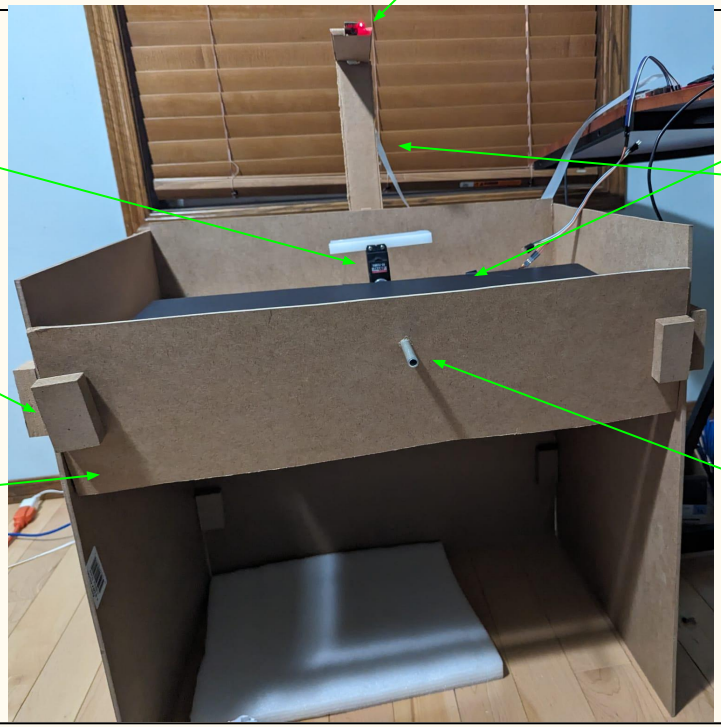
Materials

- Hs-755Mg Servo motor
- 8f by 4f MDF board
- 1.6f by 0.55f foam piece
- Metal rod
- Raspberry Pi (I used 3B+)
- Raspberry Pi camera
- HDMI cord
- Micro SD to SD converter
- Monitor (with SD port)
- Keyboard and mouse
- Hammer
- Nails and screws
- Eight 2.5in by 1.5in wood blocks
- Drill and drill bits
- Tape
- Wood glue
- 2 thumb tacks
- 24 in by 2.5in cardboard
- Jumper leads

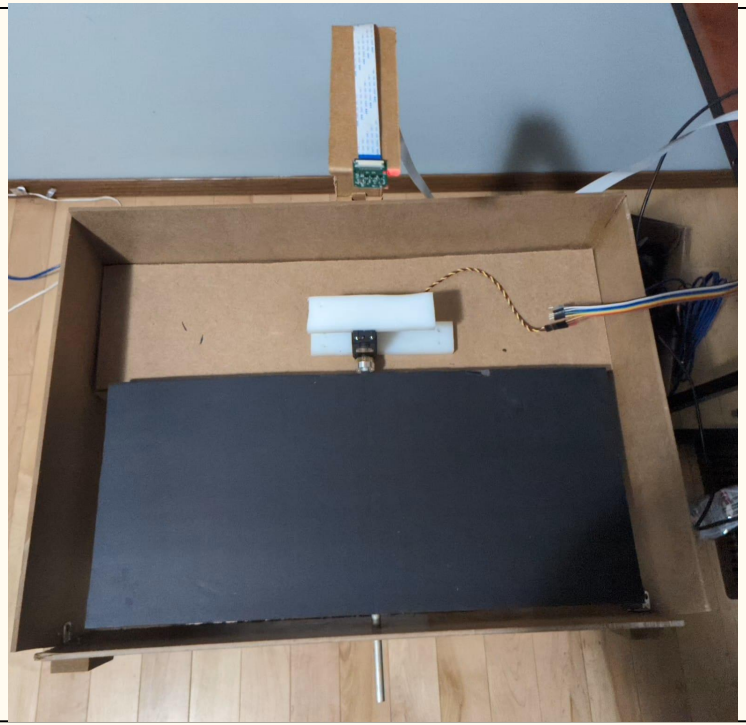
Method

1. Importing libraries and tensorflow in vs code
2. Data Importing
3. Data Visualization
4. Preparing the data
   (Transforming raw data to be read accurately when analyzed)
5. Importing OneDNN
6. Model Creation
7. Model Compilation
8. Train the Model (batch_size = 32, epochs = 10)
9. Testing Predictions
10. Save the model

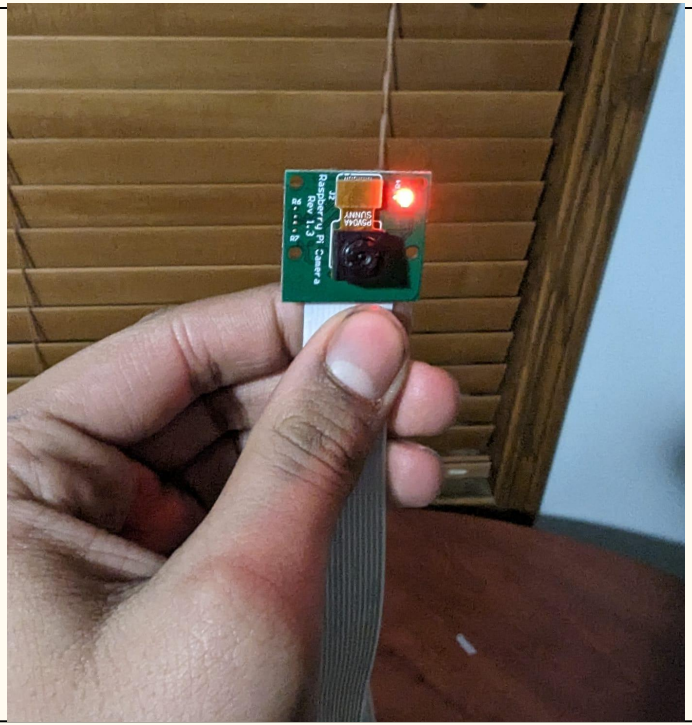
What it should look like



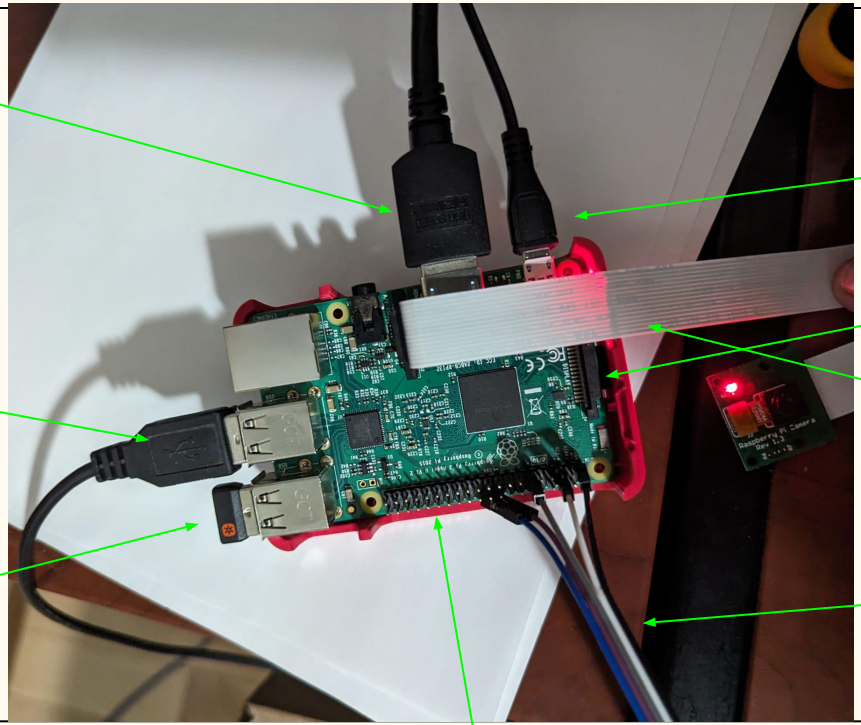
Top View



Raspberry Pi Camera



Raspberry pi 3B+



Machine learning in my project

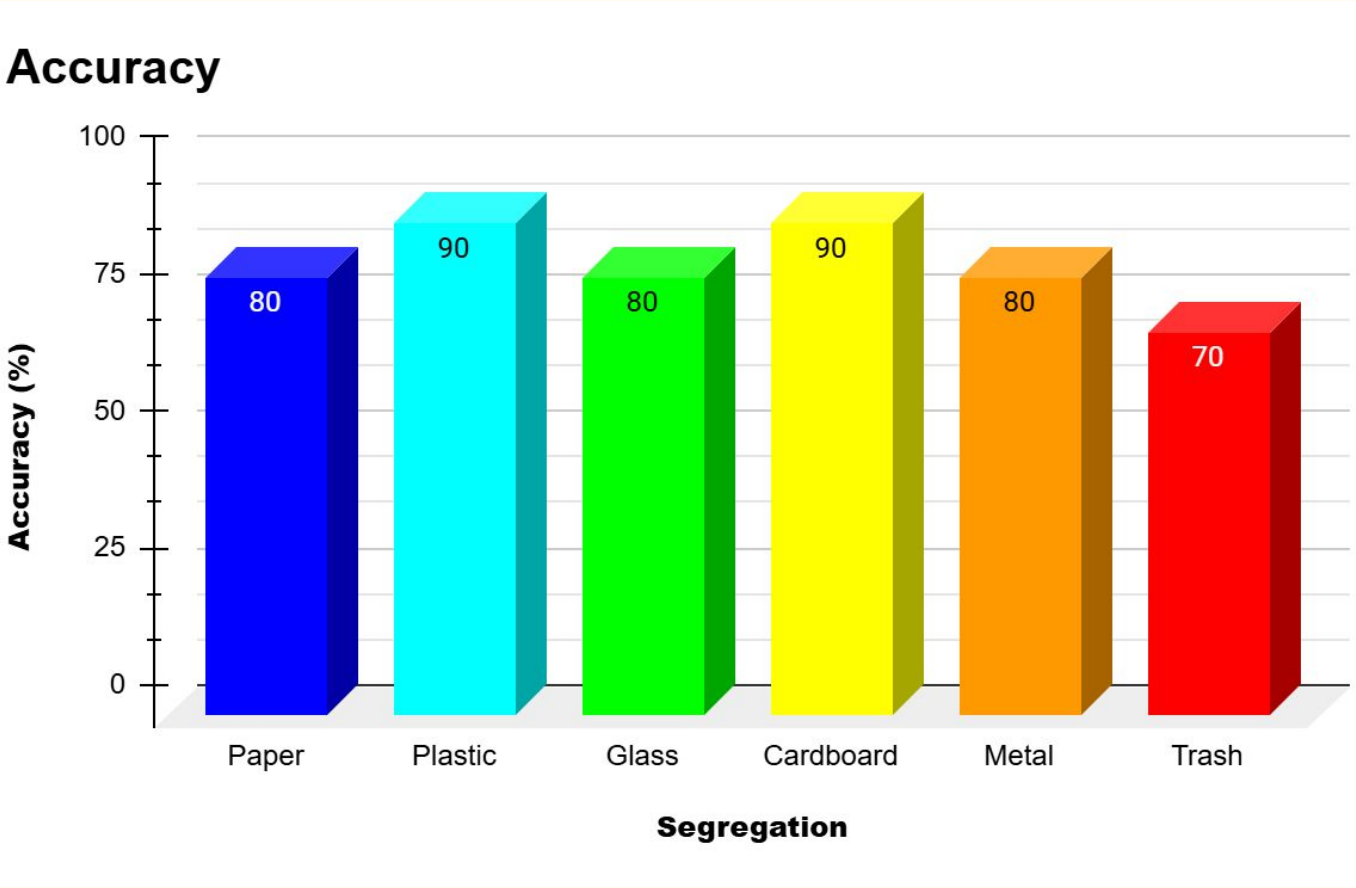
I used machine learning in the python coding in my project. Numpy is a package which is fundamental for scientific computing just for python. It is a Python library that provides a multidimensional array object, different derived objects, and an assortment of routines for fast operations on arrays. It also has mathematical aspects like logic, shape manipulation, sorting, selecting, basic linear algebra, basic statistical operations, etc. I also used PIL (Python Imaging Library) which is a library to store my Images. I also used urllib.request. This module is used to define functions making it easier in opening URLs; opening URLs is one of 2 options to upload an image onto streamlit. Tensorflow and keras are also used. Keras works with tensorflow; it simplifies the difficulties linked with deep neural networks. Tensorflow and keras will help train my model. Along with tensorflow, I am using OneDNN. It is an advanced library for deep learning that is made for optimizing the performance of deep neural network computations.

Analysis

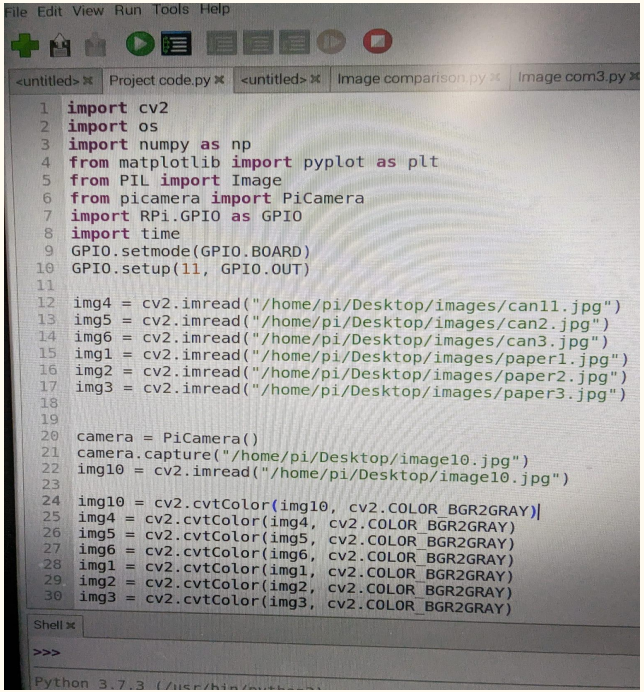
I did 10 rounds of each category to see the accuracy. Here are the results:

- Paper 8/10
- Plastic 9/10
- Glass 8/10
- Cardboard 9/10
- Metal 8/10
- Trash 7/10

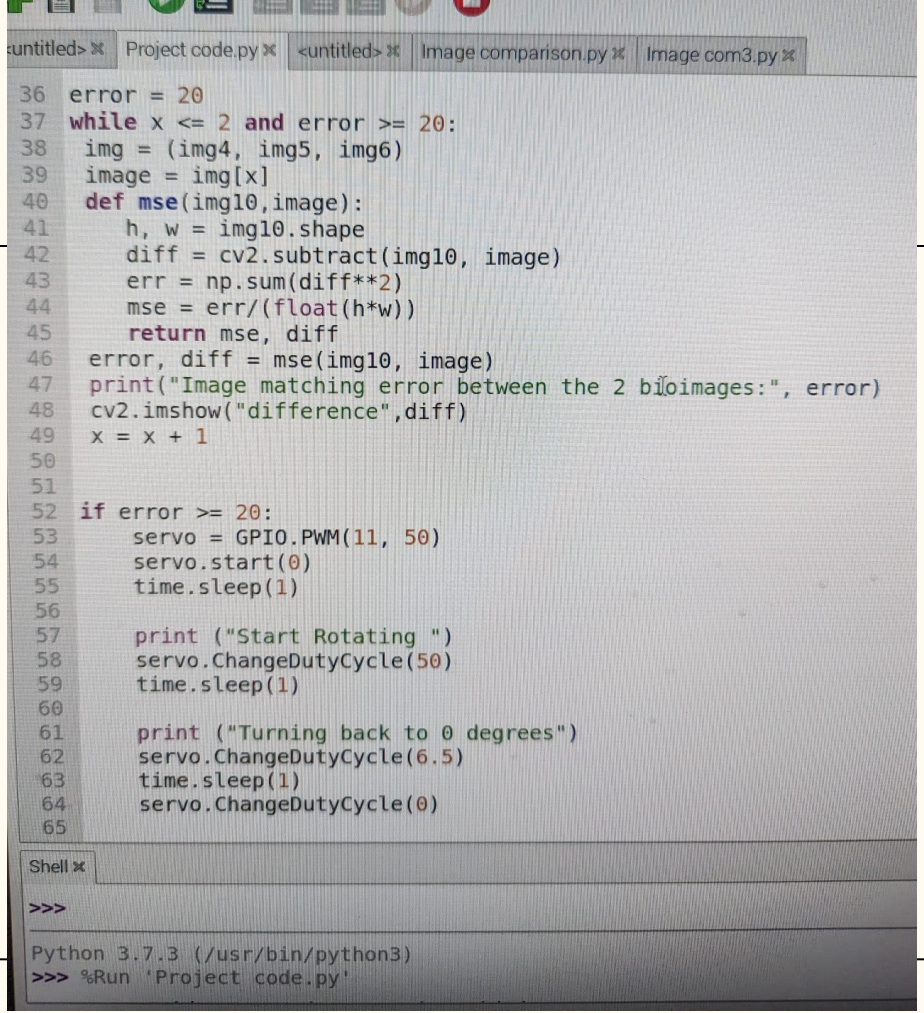
Plastic and cardboard did the best as cardboard is very similar to most images, the only time it messed up, it got confused with paper. This isn't the biggest deal though because both paper and cardboard are recyclable. Plastic has similar shapes especially when it comes to plastic bottles. Paper got mixed up with cardboard a couple of times; glass got mistaken for plastic, and metal got confused with glass and plastic. Trash did the worst because there is such a wide variety and I only have around 150 photos for it. Overall, it got a score of 49/60, Meaning in 60 rounds, it got a 78.34% accuracy.


Accuracy
100
75
50
25
0
Accuracy (%)
80
90
80
90
80
70
Paper
Plastic
Glass
Cardboard
Metal
Trash
Segregation

My code


File Edit View Run Tools Help
<untitled>
Project code.py
<untitled>
Image comparison.py
Image com3.py
1 import cv2
2 import os
3 import numpy as np
4 from matplotlib import pyplot as plt
5 from PIL import Image
6 from picamera import PiCamera
7 import RPi.GPIO as GPIO
8 import time
9 GPIO.setmode(GPIO.BOARD)
10 GPIO.setup(11, GPIO.OUT)
11
12 img4 = cv2.imread("/home/pi/Desktop/images/can11.jpg")
13 img5 = cv2.imread("/home/pi/Desktop/images/can2.jpg")
14 img6 = cv2.imread("/home/pi/Desktop/images/can3.jpg")
15 img1 = cv2.imread("/home/pi/Desktop/images/paper1.jpg")
16 img2 = cv2.imread("/home/pi/Desktop/images/paper2.jpg")
17 img3 = cv2.imread("/home/pi/Desktop/images/paper3.jpg")
18
19
20 camera = PiCamera()
21 camera.capture("/home/pi/Desktop/image10.jpg")
22 img10 = cv2.imread("/home/pi/Desktop/image10.jpg")
23
24 img10 = cv2.cvtColor(img10, cv2.COLOR_BGR2GRAY)
25 img4 = cv2.cvtColor(img4, cv2.COLOR_BGR2GRAY)
26 img5 = cv2.cvtColor(img5, cv2.COLOR_BGR2GRAY)
27 img6 = cv2.cvtColor(img6, cv2.COLOR_BGR2GRAY)
28 img1 = cv2.cvtColor(img1, cv2.COLOR_BGR2GRAY)
29 img2 = cv2.cvtColor(img2, cv2.COLOR_BGR2GRAY)
30 img3 = cv2.cvtColor(img3, cv2.COLOR_BGR2GRAY)
Shell
>>>

My code


<untitled> Project code.py <untitled> Image comparison.py Image com3.py
error = 20
while x <= 2 and error >= 20:
img = (img4, img5, img6)
image = img[x]
def mse(img10,image):
h, w = img10.shape
diff = cv2.subtract(img10, image)
err = np.sum(diff**2)
mse = err/(float(h*w))
return mse, diff
error, diff = mse(img10, image)
print("Image matching error between the 2 bioimages:", error)
cv2.imshow("difference",diff)
x = x + 1
if error >= 20:
servo = GPIO.PWM(11, 50)
servo.start(0)
time.sleep(1)
print ("Start Rotating ")
servo.ChangeDutyCycle(50)
time.sleep(1)
print ("Turning back to 0 degrees")
servo.ChangeDutyCycle(6.5)
time.sleep(1)
servo.ChangeDutyCycle(0)
Shell
>>>
Python 3.7.3 (/usr/bin/python3)
>>> %Run 'Project code.py'

My code

```
66  if error <= 20:
67      servo = GPIO.PWM(11, 50)
68      servo.start(0)
69      print ("Turning backwards")
70      time.sleep(1)
71      servo.ChangeDutyCycle(4)
72      time.sleep(1)
73      servo.ChangeDutyCycle(6.5)
74      time.sleep(1)
75
76
77  servo.stop()
78  GPIO.cleanup()
79
80
81
```

Shell ✖

Conclusion

My hypothesis is correct, my project if expanded would almost always correctly segregate trash at the source level. This will reduce the upstream sorting cost and less waste to reach landfills. So indeed this would benefit the environment, human health, and the economy. Proper waste management can also reduce pollution, prevent the spread of diseases, and conserve natural resources.

Sources of error

- It wasn't completely accurate - I have many photos in my dataset, but there will always be new angles or orientations that I do not have, therefore not making it 100% accurate

- I could have made a physical model - I made it web-based meaning I need to use images online or already saved images, it would be more efficient if it could take a picture in real-time and compare it to my dataset.

- Every time I needed to load my application, I would have to type "streamlit run app.py". After that, it would take a long time, approximately 2 minutes to load, meaning that it isn't very fast and efficient to work with.

Real World Applications

- **Would decrease the amount of unwanted trash in landfills** - So many people put the wrong trash in the wrong bins, this project makes it so even if people are lazy, everything will be segregated properly. With rates going so high to decontaminate recycling from garbage (up to $4 million), then I think that at this point, we have to try anything else. Not only is it in landfills, in countries with poverty and no access to a proper way of disposing of trash, but it all ends up on the streets, so if this project is expanded, it must be accessible for all countries. It would make a huge difference in these people's lives.
- **Cost efficiency** - Costs would go way down, there would be such a small amount of error if this project gets expanded for the trash to segregate incorrectly. Making it less decontamination, meaning also less money wasted. Even though the cost would go way down, it won't completely be reduced. There will still be some people who would be too lazy to put trash in a machine that does the work for them or they are not near a machine.
- **To show change** - If this can help the world, people will see the change many years from now and realize how much damage and harm they have done to the planet. The difference between then and now. Maybe then, people might become better people in general.
- **Relief to recycle factory people who knew how bad things were** - These people are probably one of the only people who would actually know how bad things were exactly. This would give them relief knowing that the world could be in better hands.

Bibliography - Background Research

Trash Today

Avery, R., & Specialist, M. R. (2022, October 16). 84% of UK households are unintentionally contaminating their recycling bins. WRAP. https://www.wrap.ngo/media-centre/press-releases/84-uk-households-are-unintentionally-contaminating-their-recycling-bins

Chung, E. (2018, April 9). Many Canadians are recycling wrong, and it's costing US Millions | CBC News. CBCnews. https://www.cbc.ca/news/science/recycling-contamination-1.4606893

Norman, H. (2018, April 11). Managing contaminated recyclables will cost Toronto Millions by end of year. The Globe and Mail. https://www.theglobeandmail.com/canada/toronto/article-changing-recycling-market-could-cost-toronto-92-million-this-year/#:~:text=Managing%20the%20increasing%20contamination%20will,to%20half%20of%20the%20cost.

Petricic, S. (2018, March 28). China looks to turn ban on foreign trash into growth opportunity | CBC news. CBCnews. https://www.cbc.ca/news/world/pollution-recycling-china-petricic-1.4593078

Recycling Stats

Calgary, C. (2025). Calgary Recycling Facts. https://www.calgary.ca. https://www.calgary.ca/waste/residential/recycling-facts.html#go

Zettler, M. (2019, March 27). Toronto recycling: Why so much material still goes to landfill - toronto. Global News. https://globalnews.ca/news/5099574/toronto-recycling-packaging-landfills/#:~:text=Each%20year%2C%20Toronto%20manages%20approximately,%2C%20and%20non%2Drecyclable%20materials.

What can and can't be recycled

Calgary, C. (2025a). What can go in your blue cart. https://www.calgary.ca. https://www.calgary.ca/waste/residential/what-can-go-in-blue-cart.html

Calgary, C. (2025b). What can't go in your blue cart. https://www.calgary.ca. https://www.calgary.ca/waste/residential/what-cannot-go-in-blue-cart.html

Trash Segregation problems Today

Chung, E. (2018, April 9). Many Canadians are recycling wrong, and it's costing US Millions | CBC News. CBCnews. https://www.cbc.ca/news/science/recycling-contamination-1.4606893

Bibliography - Background Research

What is Machine Learning?

Brown, S. (2021, April 21). Machine Learning, explained. MIT Sloan. https://mitsloan.mit.edu/ideas-made-to-matter/machine-learning-explained

I, bm. (2025, January 24). What is machine learning (ML)?. IBM. https://www.ibm.com/topics/machine-learning

What is python and streamlit?

mhadhbi, N. (2024, September 29). Python tutorial: Streamlit. DataCamp. https://www.datacamp.com/tutorial/streamlit

What is tensorflow?

NVI, D. (2025). What is tensorflow?. NVIDIA Data Science Glossary. https://www.nvidia.com/en-au/glossary/tensorflow/#:~:text=Heavily%20used%20by%20data%20scientists,tensors

Python, U. (2025). Urllib.request - extensible library for opening urls. Python documentation. https://docs.python.org/3/library/urllib.request.html

Bibliography - Background Research and Project

Cchangcs. (2018, November 24). *Garbage classification*. Kaggle. https://www.kaggle.com/datasets/asdasdasasdas/garbage-classification

raison024, R. (2022). *RAISON024/smart-garbage-segregation: Image classification for recycling refers to the use of machine learning algorithms to automatically classify images of waste materials, such as plastic, paper, and metal, into their respective categories.* GitHub. https://github.com/raison024/Smart-Garbage-Segregation

NumPy, T. (2020). *What is numpy?#*. What is NumPy? - NumPy v1.26 Manual. https://numpy.org/doc/stable/user/whatisnumpy.html

Python, R. (2023, January 25). *Image processing with the python pillow library*. Real Python. https://realpython.com/image-processing-with-the-python-pillow-library/

Stratvert, K. (2021, March 25). 👩‍💻 *python for beginners tutorial*. YouTube. https://www.youtube.com/watch?v=b093aqAZiPU&t=1134s