

# ML Experimental Design Discussion

## Overview of Assignment 1

Hanna Bugler  
PhD Candidate  
Biomedical Engineering  
Schulich School of Engineering

F2024

# ML Experimental Design Overview

- Problem Definition and Experimental Design Setup
- Data Collection
- Model Design and Hyperparameter Tuning
- Performance Evaluation

# ML Experimental Design Overview

- **Problem Definition and Experimental Design Setup**
- Data Collection
- Model Design and Hyperparameter Tuning
- Performance Evaluation

# Problem Definition – Assignment 1

The city of Calgary assigned your team to develop a garbage classification system that, given a cellphone picture of an object you want to throw away and a short sentence describing the object, the system tells you whether to throw it in the “green”, “blue”, “black” trash bin or somewhere else. You can see more information about the city of Calgary's garbage collection system here: <https://www.calgary.ca/uep/wrs/what-goes-where/default.html>.



Whenever we're starting with a new ML problem, a literature review of similar works can help provide a good starting point!

# Experimental Design Considerations – Assignment 1



Data

Source, Diversity, Preprocessing

# Assignment 1 – Data Source

- Write a quick description of the following piece of garbage which would help with sorting it?



# Assignment 1 – Data Diversity

- How does garbage vary?
- How do images vary?
- How may these effects bias the model?





# Assignment 1 – Data Preprocessing

- Object segmentation
- White background
- Pre-filtering
- Convert images to grayscale
- The object material alone does not determine the appropriate trash bin





# Experimental Design Considerations – Assignment 1



Task Definition

Architecture, Output type, Metrics

# Assignment 1 – Task Definition

## Model Task complexity



Multi-object detection

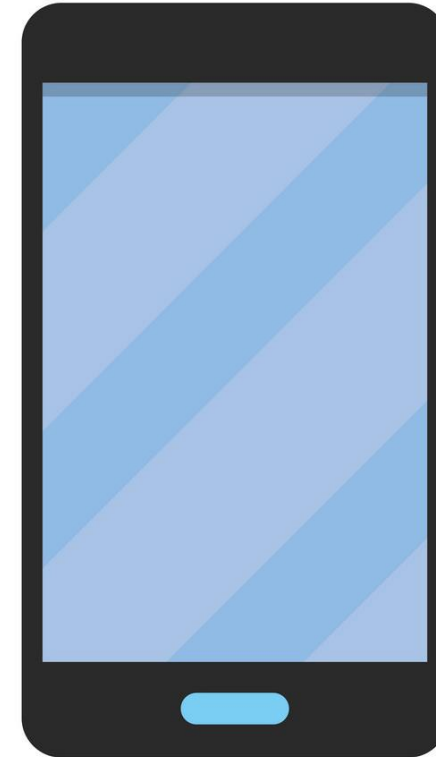


Single object background segmentation

**Pros vs. Cons?**

# Assignment 1 – Task Definition

**Hardware limitations?**



**"Need a lighter high-performing model because of the computational limitations on cell phones when a new sample is detected."**

# Experimental Design Considerations – Assignment 1

Addressing  
Anticipated Issues

Generalizability, Anticipated Model Uses

# Assignment 1 – Potential Issues

**Rare items that your model did not see during training?**



**"Another issue that can arise is the inability of the system to classify certain objects due to their rarity and the lack of data. A prime example would be Floppy Disks or VHS tapes that are generally not sold commercially anymore and are hard to find."**



# Assignment 1 – Potential Issues

Not green, nor black nor blue trash bin?





# Assignment 1 – Potential Issues

**Green or blue? What if the box is closed?**



**“Another potential issue is the class overlap for certain items. For example, a clean pizza box can be disposed of in the blue bin, however if it is greasy/covered in food toppings, it should be disposed of in the green bin. Depending on how the photo is taken, it may be impossible to know for sure which class this item belongs to.”**

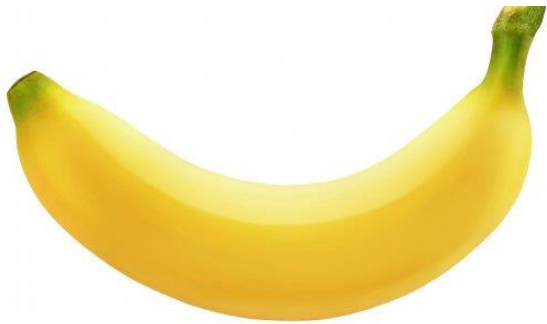


# Assignment 1 – Good Suggestions

- Crowdsourcing and future data collection
- “Once the photos will be taken from different cameras with different resolutions, we have planned to train and validate the model with some cell-phones and test them on the others to check our model's generalization.”

# Assignment 01 – Philosophical question...

**Trash?**



# Thank you!



UNIVERSITY OF  
CALGARY