MAIS 202 - Project Proposal:
Waste Classification to Increase Recycling and Compost Quality

Improper residential and commercial waste sorting is a serious problem for waste management services and climate change mitigation. Improper waste sorting can lead to the "contamination" of an entire batch of recycling. For example, if a half-full coffee cup is thrown into a bag of paper recycling, that entire load of recycling is rendered unusable, and waste management services are left with no choice but to throw out the entire load. Furthermore, when compostable items such as food waste and food-soiled cardboard boxes are thrown into landfill, they anaerobically decompose at landfill sites [1], releasing methane gas (a greenhouse gas which traps 25 times more heat than carbon dioxide [2]). In 2022, emissions from landfills accounted for 17% of Canada's national methane emissions. [3].

McGill has made efforts to improve the clarity and readability of waste sorting posters around campus, but still students and staff routinely dismiss these posters, or fail to understand certain nuances around waste sorting that these posters fail to capture. This project aims to train, test, and deploy a CNN to identify waste items into their appropriate waste streams, creating an educational McGill-specific tool for proper on-campus waste sorting practices.

I. Choice of Dataset

https://www.kaggle.com/datasets/alistairking/recyclable-and-household-waste-classification

McGill's nonhazardous waste is split into 4 streams: *Compost*, *Paper/Cardboard*, *Recyclable containers*, and *Landfill*. The chosen dataset classifies images into a variety of categories that each fall into exactly one of McGill's 4 waste streams. For example, the dataset has a collection of images classified as plastic shoppings bags (which would go into *Landfill*; thin plastics are not profitable enough to recycle), and a collection classified as Plastic soda bottles (which could go into *Recyclable containers*). Other datasets group all plastics into one classification, but this high-quality annotation allows for more nuance in the model's results.

II. Methodology

A. Data Preprocessing

As described above, the images are pre-classified into common waste items/categories, which can later be sorted into one of McGill's 4 nonhazardous waste streams. The bulk of the data preprocessing in this project would pertain to making the images diverse and efficient for training the CNN. Each image will be rescaled and padded to a common dimension (padding should not impact the behaviour of the CNN)[4]. To make training more efficient, pixel values will be normalized to a range from 0 to 1 ([0,1]). To increase the diversity of the dataset, image rotation and flipping will be applied.

B. Machine learning model

A Convolutional Neural Network (CNN) will be trained to classify waste items. CNNs are commonly used in image classification, and well-adapted to handle the type of dataset chosen for this project. Residual Networks (ResNets) are a type of CNN now more commonly used for their ability to use deeper layers while avoiding making the model's gradient too small/large. ResNet will be considered the model of choice for this project.

C. Evaluation Metric

Accuracy, precision, and recall will be used to measure the model's ability to classify waste items. This project aims to attain a waste classification accuracy of at least 75%.

III. Application

The webapp portion of this project will be the most useful for highlighting common waste sorting errors to the user.

Users of the webapp will be able to upload an image of a waste item. The webapp will process and input submitted images to the model, which will classify the image into one of the chosen dataset's categories. Using the model's result, the webapp will identify the item as being a part of one of McGill's 4 nonhazardous waste streams. Certain nuances in waste sorting will be represented in the webapp. For example, one nuance which will be represented is that recyclable items can only be recycled if they contain no food residue. For instance, a Type 1 plastic food container (which is recyclable if clean) that has an oily or saucy residue inside it should be thrown into landfill. Another example is that a cardboard pizza box soiled with oil should be thrown into compost ("brown" paper items like cardboard and paper napkins should go into compost when soiled with food). Using the CNN to visually recognize whether a recyclable item has been soiled with food is outside the scope of this project, so instead a prompt will be given to the user, asking if the item has food residue for example, resulting in the webapp's final classification of the item.

Citations

[1]https://www.canada.ca/en/environment-climate-change/services/managing-reducing-waste/municipal-solid/waste-greenhouse-gases-canada-actions.html
[2]https://www.canada.ca/en/environment-climate-change/services/managing-reducing-waste/municipal-solid/waste-greenhouse-gases-canada-actions.html
[3]https://www.canada.ca/en/environment-climate-change/services/climate-change/greenhouse-gas-emissions/inventory.html
[4]https://journalofbigdata.springeropen.com/articles/10.1186/s40537-019-0263-7