Introduction:

Assignment 3 of data collection for the use of object detection using Yolo V8 was based on collecting and training data set for image detection. Our team focused on items such as Metal lid (from jar or bottle), Non-prescription medication bottle (empty), Plastic produce containers,

Plastic single-serve, Plastic straw, Plastic sunscreen bottle, Plastic cups (single serving size) & plastic pots & trays.

Majority of the images were collected from primary sources that is our individual phones and secondarily from google, friends and families. While collecting data we made sure that the images have different background noises and multiple lights on the pictures to train the model. Our target was to source minimum 120 images for each group and rename, label the images using LabelBox and then save excel file with the metadata containing properties of images like phone model, format & the dimensions. The result would be to train the model so that it can detect pictures with a 90% precision. We used python code to rename the image files and export the data to CSV and manually added the phone model.

Image collection:

Images for each category was collected using both primary (mobile devices) & secondary source (google images and other websites). For primary method we used both android and Iphone devices to gather images in JPEG format. The lighting condition was taken into consideration along with different background. For instance, images were taken in broad daylight to get the sun exposure with background noises while some were taken in night with and without streetlights (with flash). The night light helped us to gather some very dark images which are hard to understand by machine resulting in better training for our model.Along with images captured under sodium vapour lamp (streetlights), few images where captured using night lamps . These pictures are very hard to understand and hence has very broaden axis when we did the labelling using makesense.ai to set the labels for image annotation. We used outside real scenarios wherein the pictures were captured from garbage and in hand to get the real-life scenarios for image detection. These pictures will be a great way to train the model when we will be working on our team project for these pictures.

Secondary pictures included images from web (Google), family members and friend’s phone. These were from a variety of sources and devices including both Android whereas primary pictures included Android devices only. While getting these pictures we used WhatsApp to transfer the pictures from our families & friends. These might have tampered with the actual resolution and hence it dropped for few pictures like from Iphone to 900x1600. While a few were taken from google drive of us and had many different resolutions as we had multiple mobile phones till now. These pictures were combined category wise and then stored on drive to rename & label.

Challenges:

During image collection we did not many pictures to reach 120 mark and hence we used secondary sources like Google and asked family members and friends. But later, we deleted multiple google images and searched/captured pictures and replaced these google images with us. For background, to make it tougher we used multiple conditions like keeping it with multiple color objects, lighting conditions, different source location like in malls, trolleys, baskets, parks, grass, dustbins, and different timelines like in morning, afternoon, evening, and night along with bright light like in Sun, in night under the streetlamps and in house during night under the study lamps. These pictures are a collection of different lighting conditions and phones.

Some of the images were hard to find because it did not have multiple types like the color was constant and any other was not available like for cups (majority were red). For pots & trays, trays were a bit difficult to find even though we visited a lot of stores. For plastic straws, the majority was taken from stores and was hard to create different stores as we did not have the same at home. Similarly, for sunscreen, the variety was less and hence we had to create multiple scenarios at home.

Most importantly, we had stir stick as a category to collect and train data, but after consideration we came to the conclusion that it’s difficult to find and collect 120 images for this category and hence, we replaced it with plant pots and trays.

Finally, to summarize, image collection and labelling was done manually and consumed a lot of time to get the required counts for each category while image renaming and exporting file to csv to create metadata was achieved using python code. Even in the metadata, phone details along with android version it was manual work to get the details of each image and store it in a new column with details for metadata.