Ralfy Chettiar
Report (Assignment)

Objective:

As per the given problem statement we were supposed to identify type of recyclable bottles and can out of the 7 types. Samples of the bottles on the conveyor were given and based on those images a machine learning model had to be made and trained and predict the objects. To achieve that result I followed few steps which are mentioned below.

Execution:

Step 1: - *Identify each object from a single image.*

- To achieve the above result, I first had to separate each object from the background as foreground. The given rgb (png) images were loaded and stored in an array. Then the semantic segmentation of the images was loaded which were in the form of.exr. For the code to run those files they had to be converted into readable format which in this case were png's. Once the conversion was done, they too were stored in an array.
- Then in a function each image was converted to a grayscale image. The masked regions were labelled to identify each component so that based on the labels they can be identified and further processed.
- The labelled images were used to find their area, extent, intensity, etc and stored into a dataframe from pandas.
- Using the labeled images, the extent of each component was found and that was then used to get separate the bottles/cans from the color image. These images were stored in a separate array so that they can be used to train the convolution neural network.

Step 2: Creating the dataset for the neural network

- Once the objects from all the 200 sample images are separated, they are then segregated into their respective categories.
- There are 7 categories in total and those images are identified for their type and stored in individual folders.
- For recognizing the objects, I used color to create a mask and then calculate the area which is detected. Based on the set threshold the function then stores them in their respective folders.
- These folders are then used to create labels and the dataset for the modeling is obtained. The inbuilt keras function is used to create the sets and assign labels to them.

- After the data is created then a CNN network is developed which learns from the data and then makes predictions.

Step3: Convolutional Neural Network

- In this the images are sent through a series of filters to extract the max number of features, and this helps the model to learn the unique characteristics of the object.
- The filters are divided into number of layers. Conv2d, normalization, maxpooling and relu as an activation function was used.
- At the end of the fully connected network SoftMax was used to identify the labels and predict the results.
- Depending on the number of epochs and the batch size the accuracy of the model varied.

Step 4: Tuning

- Here by changing the learning rate of the optimizer (Adam in this case) we can vary the results rate.
- The arrangement of the filters can help extract features which are specific to that type of object.

Result:

Using the developed model, I was able to get an accuracy of about 90% and the number of epochs used were 50. The prediction can be made stable further by tuning the parameters. Also, the dataset generated can further be improved so that only correct data is provided and the model trains on that.

Improvements:

If I had some more time, I would have been able to pre-process the data better by changing the threshold. For the neural network the arrangement of the filters would have helped get a smoother curve for the training. Changing these and a minor adjustment in the thresholds we could improve the model prediction.

Completion:

With the given time frame, I was able to build a neural network that took in the train data set and labels and predict the results. However, I was not able to find the orientation of the bottles/cans, spatial boundaries, and labels. These things could have been achieved by me by the below steps.

I. To detect the orientation of the objects I could use Principal Component Analysis (PCA) to detect the majority of the points and accordingly draw a line which takes in max number of points and using the base line (taking the image x-y coordinates as reference) I could calculate the angle between them and then could have displayed the line on to the image.

II. For outputting the spatial boundaries, OpenCV's draw contours can be used to detect the objects by taking the cropped images as reference and then by specifying the parameters in the function could have drawn a rectangle around each object.

III. The OpenCV function helps to put in text as well along with the box. Using this the labels of each object can be printed at any desired location around the box.

IV. This can be done to each image containing 8 bottle and based on the prediction outputted the results on the main image and saved in a separate folder.