Section III: Results

After finetuning the CLIP zero-shot classification model, the accuracy of the model was tested using prepared datasets. Using the text encoders “Organic Food Compostable” and “Plastic Glass Metal Cardboard Recyclable,” a set of testing images was run through the model to test its accuracy. The accuracy for both waste categories, organic compost, and plastic recyclables, is relatively high considering that the model was not finetuned with thousands of images, with the basic accuracies of 92% and 75% respectively.

Along with the classification accuracy rates for each waste category, a confusion matrix was constructed to gain further insight into where the model was having difficulty correctly identifying the waste. The finetuned model classified 458 images that were organic compost to be organic compost but classified 42 images that were actually organic compost to be plastic recyclables. For images that were plastic recyclables, 373 of them were accurately classified as plastic recyclables but 127 of them were misclassified as organic compost.

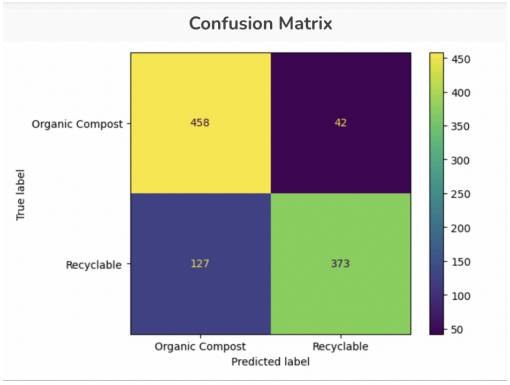

Figure 4: Confusion matrix for the CLIP zero-shot classification model prior to finetuning.

F1 Score

	precision	recall	f1-score	support
Organic Compost	0.95	0.74	0.83	1401
Plastic Recyclable	0.74	0.95	0.83	1112
accuracy			0.83	2513
macro avg	0.84	0.84	0.83	2513
weighted avg	0.86	0.83	0.83	2513

Figure 5: F1 score for the CLIP zero-shot classification model prior to finetuning.

The F1 scores for both waste categories were also calculated using the precision and recall values. The F1 score is the harmonic mean of the precision and recall values, meaning that it is two times the product of the precision and recall values divided by the sum of the precision and

recall values. The F1 is an important statistical test to run as it gives insight into the trade-off between precision and recall values. Therefore, F1 can be used to measure how effectively the model makes that trade-off. The F1 score displayed for the organic compost and plastic recyclable categories is close in value since both their precision and recall values are also relatively close in value as well.