############ MODEL Number 1 ########################

The confusion matrix shows how the model performs across all classes:

**Best Performing Classes:**

Food_Waste: 232 correct predictions (relatively few misclassifications)

Plastic: 478 correct predictions (though with many misclassifications into other categories)

**Problematic Classes:**

Glass: Only 85 correct predictions with 106 misclassified as Plastic (major confusion)

Metal: 123 correct but significant misclassifications as Plastic (82)

Other: Only 72 correct with many misclassified as Plastic (110)

Paper: 183 correct but notable misclassifications as Plastic (60)

**Common Misclassification Pattern:**

Many classes are frequently misclassified as "Plastic", suggesting the model may be biased toward this category or that plastic waste shares visual features with other waste types.

**Accuracy Metrics**

The model achieved a training accuracy of 0.9 (90%), which appears good at first glance.

However, the confusion matrix reveals significant class imbalance issues and poor performance on several categories despite the high overall accuracy.

**Key Findings**

Class Imbalance Issues:

The number of correct predictions varies greatly (from 72 for "Other" to 478 for "Plastic"), suggesting uneven class representation in the dataset.

**Model Biases:**

The model shows a strong tendency to classify ambiguous cases as "Plastic", which could indicate either: Plastic waste is overrepresented in training data Plastic waste shares visual features with other categories The model hasn't learned distinctive enough features for other classes

**Particular Challenges:**

Glass waste seems particularly difficult to classify correctly (only 85 correct vs 106 as Plastic) The "Other" category performs poorly, which is expected but may need better definitio

**Recommendations for Improvement**

Data-Level Improvements:

Balance the dataset across all categories Review labeling consistency, especially for problematic classes Consider merging or redefining categories that are consistently confused (e.g., Glass and Plastic)

Model-Level Improvements:

Implement class weighting to handle imbalance Try different architectures or add attention mechanisms Apply more aggressive data augmentation for minority classes

################################### MODEL Number 2
###################################

################################### ResNet 18
###########################################

** Accuracy: 87%**

This is the overall correct prediction rate on the validation set.

**Strongest Classes:**

Metal: - Excellent on both precision (0.90) and recall (0.91). Glass: - High precision (0.93), though slightly lower recall (0.84). Plastic: - Good recall (0.91) — important since it has the most support (472 samples).

**Slightly Weaker Classes:**

Paper and Other have the lowest recall (~0.79) — the model misses some true cases here. Other and Cardboard are also slightly confused with each other in the confusion matrix.

**Macro & Weighted Averages:**

Macro Avg F1 = 0.87 → Equal weight to each class — suggests balanced performance.

Weighted Avg F1 = 0.87 → Also good, means model handles dominant classes (like Plastic) well.

**Conclusions**

ResNet-18 model is performing very well with an overall 87% accuracy and strong per-class F1 scores. It's suitable for deployment in most practical waste classification setups, with scope for minor improvement via tuning or data enhancements.