

# Recycling Dataset

W281 Final Project



# Proposal | Automate sorting of recyclables



**Batteries**



**Plates**



**Paper**

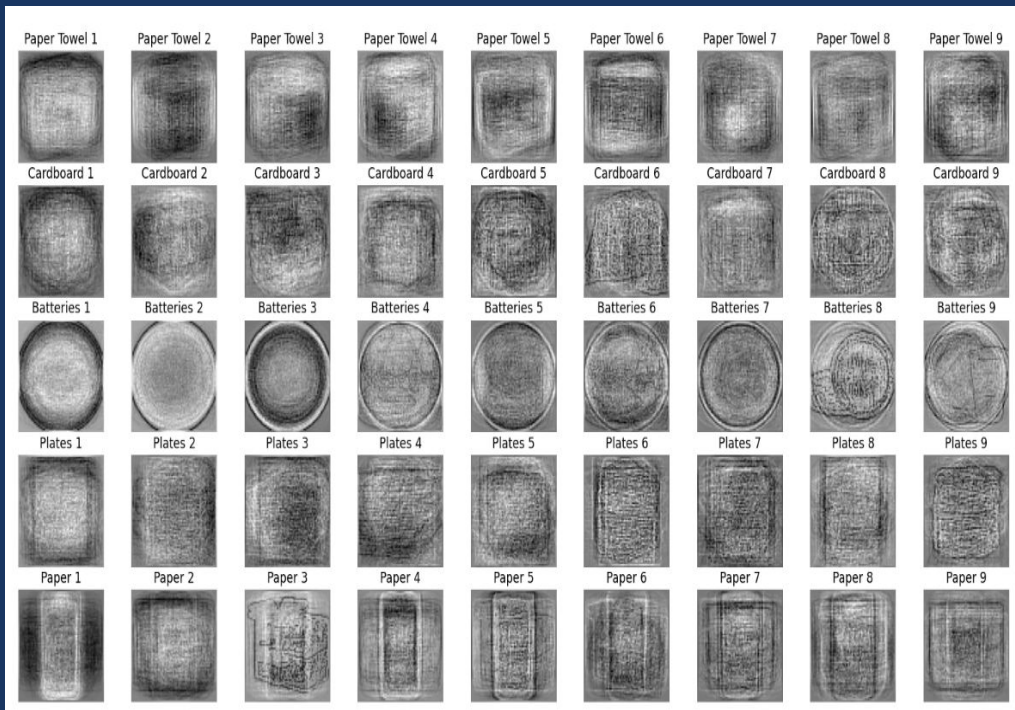


**Paper  
Towels**



**Cardboard**

# Features



1. Edge detection
2. Texture detection
3. Color
4. Texture
5. Key point detection

# Classification | Approach



# Classification | Baselines

## Dummy Classifier

---

- ❖ Most Frequent Class
- ❖ 22% Accuracy

## Baseline SVM Classifier

---

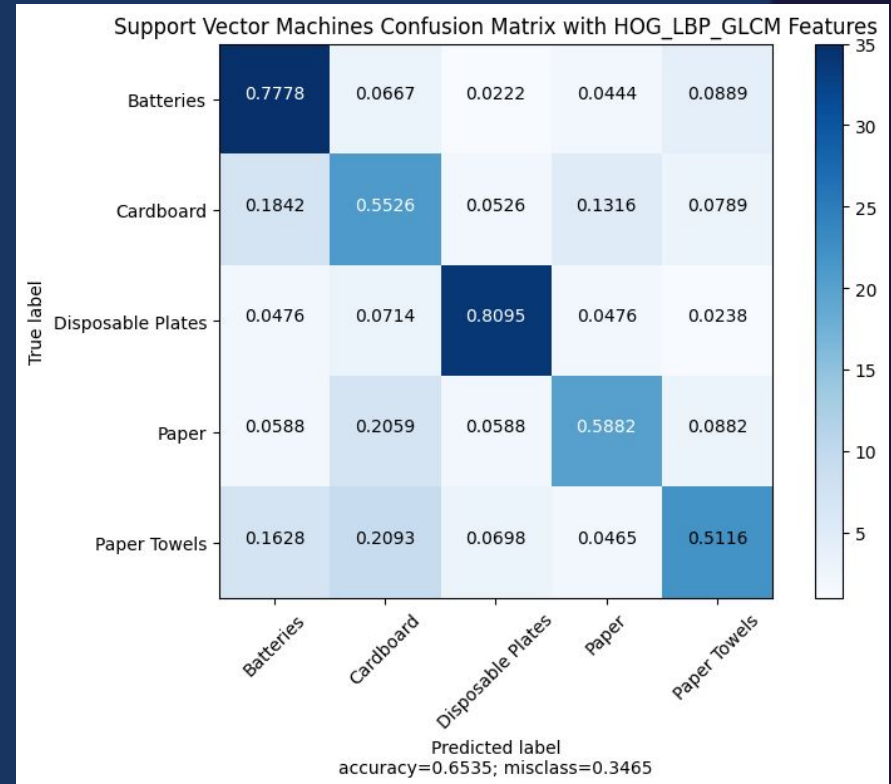
- ❖ Gray scale, resized, without background
- ❖ No Hyperparameter tuning
- ❖ 52.7% Accuracy

# Classification | Model Creation

- ❖ Applied same preprocessing to all image
  - Fit : Training Data
  - Transform: Validation/Test Data
- ❖ Train/Validation/Test Split
- ❖ PCA
- ❖ GridSearchCV
- ❖ K-Fold Validation
- ❖ Hyperparameter tuning of features

# Classification | Best Model

- ❖ SVM – 5 Folds
  - 60.8 % Average
  - 65.4 % Max
- ❖ Batteries + Paper Plates performed best



# Efficiency | Accuracy

Model Type	Training Time	Prediction Time	Accuracy
Logistic Regression	14.1 $\mu$ s	16.2 $\mu$ s	37.4%
Tensorflow Basic Model	17.9 $\mu$ s	4.05 $\mu$ s	27.0%
SVM	2.86 $\mu$ s	3.1 $\mu$ s	60.8%



# Generalizability

## Pros

---

- ❖ A lot of variety of data within classes (not-clean)
- ❖ KFold/Cross Validation technique used
- ❖ Multiple features to get a holistic view
- ❖ Train/validation/test

## Cons

---

- ❖ Small dataset
- ❖ Much more variety in the real world
- ❖ Artificially created images

Thank you | Questions?