# Technical Report – Image Classification Project

| ID | Name |
|---|---|
| 20220250 | Kareem Mohamed Hasseeb |
| 20220191 | Abdelrahman Saeed Mohamed |
| 20220054 | Osama Mohamed Al Amrawy |
| 20220503 | Yousef Medhat Ahmed |

# 1. Introduction

The goal of this project is to develop an image classification pipeline capable of recognizing six waste categories: cardboard, glass, metal, paper, plastic, and trash. The pipeline integrates **data augmentation, deep feature extraction using MobileNetV2, and classical machine learning classifiers (SVM and KNN)**. This report summarizes the methods used, performance evaluation, and compares classifier architectures and feature extraction methods.

# 2. Dataset and Data Augmentation

The original dataset contains images with the following distribution:

| Class | Original Count |
|---|---|
| Cardboard | 259 |
| Glass | 401 |
| Metal | 328 |
| Paper | 476 |
| Plastic | 386 |
| Trash | 110 |

The dataset was split into **train and validation sets using an 80:20 ratio**.

To improve generalization and meet the assignment requirement of **at least 30% increase in training samples**, the training set was augmented to **500 images per class** using the following techniques:

- **Rotation:** small rotations between ±15° to simulate different object orientations
- **Scaling:** 0.9–1.1 zoom to account for size variations
- **Horizontal flipping:** 50% probability to mimic mirrored views
- **Brightness and contrast adjustment:** 0.8–1.2 to simulate different lighting conditions

These augmentation techniques introduce realistic variations in the training data, improving model robustness to changes in orientation, scale, and lighting, and reducing overfitting.

## 3. Feature Extraction

### 3.1 MobileNetV2 Features

- A pretrained MobileNetV2 model (ImageNet weights) was used for feature extraction.
- The top classification layer was removed, and global average pooling was applied, producing 1280-dimensional feature vectors.
- **Advantages:**
  - Captures high-level semantic features effectively
  - Lightweight and computationally efficient
  - Robust to variations introduced via augmentation

### 3.2 Preprocessing

- Images resized to 224×224 and converted to RGB
- Normalization using preprocess_input
- Features standardized using StandardScaler
- **PCA** applied to reduce dimensionality while retaining 95% variance, improving classifier efficiency and removing redundant features

The combination of deep features and PCA ensured high-quality feature representation with reduced computational overhead.

## 4. Classifier Architectures

### 4.1 Support Vector Machine (SVM)

- **Kernel:** RBF
- **C:** 10
- **Gamma:** scale
- **Probability estimates enabled** to support rejection thresholding

**Advantages:**

- Effective in high-dimensional feature spaces (1280-dimensional MobileNetV2 features)
- Handles non-linear boundaries
- Works well with limited training data
- Confident predictions can be separated from uncertain inputs via a threshold

**4.2 K-Nearest Neighbors (KNN)**

- **Neighbors:** 3
- **Distance metric:** Euclidean
- **Weights:** distance-based
- Rejection applied for low-confidence predictions (threshold = 0.6)

**Advantages:**

- Simple and interpretable
- Non-parametric, adapting to complex decision boundaries
- Effective with sufficient training data

## 5. Architecture Comparison and Trade-offs

| Aspect | SVM | KNN |
|---|---|---|
| Accuracy | 90% | 82% |
| Computational Cost | Training is slower, inference fast | Training is fast (store features), inference slower with large data |
| Scalability | Handles high-dimensional features well | Sensitive to large datasets due to distance computation |
| Robustness | Better generalization with RBF kernel | Sensitive to outliers and feature scaling |
| Handling Unknown Inputs | Rejection via confidence threshold | Rejection via confidence threshold |

**Observations:**

- **SVM** performed better due to its ability to handle high-dimensional, non-linear feature spaces.
- **KNN** is simpler but slower at inference for large datasets and slightly less precise.
- Both classifiers benefit from MobileNetV2 features, which capture semantic information more effectively than raw pixels or traditional handcrafted features (e.g., HOG, SIFT).

**Feature Extraction Comparison:** Using deep CNN features (MobileNetV2) provided better generalization and higher accuracy than traditional handcrafted features would, especially when combined with data augmentation.

## 6. Performance Evaluation

| Classifier | Accuracy | Notes |
|---|---|---|
| SVM | 90% | Robust, high precision, effective with deep features |
| KNN | 82% | Simple, sensitive to feature scaling, lower accuracy |

Coverage represents the fraction of predictions above the rejection threshold. Thresholding ensures uncertain or potentially unknown inputs are labeled as “Unknown,” increasing robustness in deployment.

## 7. Conclusion

The project demonstrates that **deep feature extraction using MobileNetV2** combined with classical ML classifiers can achieve high accuracy for waste image classification.

- **SVM** achieved higher accuracy and better robustness than KNN.
- **KNN** is a simpler alternative, suitable for smaller datasets or rapid prototyping.
- **Data augmentation** improved model generalization, making the classifiers robust to variations in orientation, scale, and lighting.

- Using **MobileNetV2 features** is highly effective compared to raw pixel data or handcrafted features.