Multi-Level Waste Classification System

Group 14 Members

Khin Yadanar Hlaing: khinyadanarhlaing.kt@gmail.com

Myo Myat Htun: myomyatskywalker@gmail.com

Aye Nandar Bo: ayenandarbo24@gmail.com

1. Literature Review

Introduction

Improper waste management is a growing environmental and social concern that contributes to pollution, greenhouse gas emissions, and resource depletion. Automated waste classification using artificial intelligence has emerged as a promising solution to reduce manual labor and increase recycling efficiency. Reviewing existing literature in this domain is essential to identify proven methodologies, performance benchmarks, and research gaps that can guide the development of our proposed waste classification system.

Summary and Synthesis

Two major research works serve as references for our project:

Paper 1: Trash Detection — Advanced Classification of Waste Materials Using ML Techniques (Khetarpal & Khetarpal, 2024)

This study developed a **Convolutional Neural Network (CNN)** model for classifying five waste categories: paper, plastic, metal, glass, and cardboard. The authors performed image preprocessing (resizing, normalization, augmentation) and used a dataset of over 25,000 labeled waste images. The model achieved **93.39% accuracy**, demonstrating that CNNs can effectively extract spatial and texture features from waste images. The study highlighted how deep learning can replace manual waste sorting and enhance recycling efficiency, but it lacked hierarchical classification (e.g., recyclable vs. non-recyclable levels).

Paper 2: An Automated Waste Classification System Using Deep Learning Techniques — Toward Efficient Waste Recycling and Environmental Sustainability (Knowledge-Based Systems, 2025)

This research proposed a **three-stage deep learning architecture** using the large-scale **TriCascade dataset (35,000+ images)**. The system implemented lightweight CNN models and achieved **96% accuracy** in multi-class classification, also supporting real-time sorting through a hardware prototype. The integration of computer vision with IoT hardware demonstrated its industrial scalability and sustainability impact. However, the study noted challenges in imbalanced datasets and model generalization across diverse lighting and background conditions.

Comparison and Insights:

Both papers confirm the effectiveness of CNN-based models for waste classification. The first emphasizes model accuracy in basic waste categorization, while the second extends toward real-time application and scalability. However, neither study explores a hierarchical classification structure (recyclable vs. non-recyclable → waste type → hazardous class), which is the innovation our project aims to introduce. Our approach also focuses on integrating multiple public datasets for improved generalization and robustness.

Conclusion

The reviewed literature shows strong evidence that deep learning techniques, especially CNNs, can significantly improve the efficiency and accuracy of waste classification. Yet, current studies often handle limited class scopes or lack hierarchical sorting. Our proposed project contributes to filling this gap by implementing a **multi-level classification architecture** capable of simultaneously identifying recyclability, waste type, and hazardous properties. This enhancement aligns with sustainable development goals related to waste reduction, health, and innovation.

2. Data Research

Introduction

Reliable and diverse data is essential for training a deep learning model capable of handling various waste materials under real-world conditions. Exploring and selecting appropriate datasets ensures that the model learns to generalize across different lighting, angles, and material textures, leading to more robust classification performance.

Data Description

We identified and reviewed several publicly available datasets for waste image classification:

1. **Waste Segregation Image Dataset (Kaggle)**

    - **Source:** Kaggle Dataset
    - **Format:** JPEG/PNG images organized into folders for recyclable and non-recyclable waste.
    - **Size:** ~12,000 images.
    - **Relevance:** Suitable for binary classification (recyclable vs. non-recyclable).

2. **Waste Classification Dataset (Kaggle)**
    - **Source:** Kaggle Dataset
    - **Format:** 2D image data with clear labeling into six waste types: paper, plastic, glass, metal, cardboard, and trash.
    - **Size:** ~25,000 images.
    - **Relevance:** Ideal for multi-class classification tasks based on material type.

3. **Garbage Classification Dataset (Kaggle)**
    - **Source:** Kaggle Dataset
    - **Format:** JPEG images categorized into 12 waste classes.
    - **Size:** ~20,000 images.

    Relevance: Provides fine-grained labels for complex waste classification.

4. **Roboflow Waste Datasets**
    - **Source:** [Roboflow Universe]
    - **Format:** Annotated image datasets with bounding boxes for object detection tasks.
    - **Relevance:** Useful for extending the model to real-time detection and localization.

Data Analysis and Insights

The selected datasets together provide a rich variety of waste categories, textures, and environments. The combination of these sources will allow the creation of a **balanced and comprehensive dataset** for training and testing. Preprocessing steps such as **image resizing (224×224)**, **normalization**, and **augmentation** (rotation, flipping, brightness adjustment) will improve model generalization.

An exploratory data analysis (EDA) phase will assess **class distribution**, **color variation**, and **texture differences** to identify imbalance or bias in the dataset, and oversampling or augmentation (like SMOTE for image embeddings) will be used to address class imbalance.

Conclusion

A thorough dataset exploration confirms that multiple open-source datasets can be merged and preprocessed to create a diverse and representative dataset for hierarchical waste classification. The resulting dataset will support three classification levels (recyclable vs. non-recyclable, waste type, hazardous class) to ensure comprehensive environmental applicability and strong model performance.

3. Technology Review

Introduction

A technology review is crucial to identify the tools, frameworks, and algorithms that will enable efficient model training, deployment, and performance optimization. The technologies selected must align with the project's objectives: real-time image classification, scalability, and interpretability.

Technology Overview

The main technologies and tools for this project include:

1. **Convolutional Neural Networks (CNNs):**
    - Core deep learning architecture for image classification.
    - Extracts spatial and texture features using convolutional and pooling layers.

2. **Transfer Learning (MobileNetV2, ResNet50):**
    - Utilizes pre-trained models on ImageNet to reduce training time and improve performance with limited data.
    - MobileNetV2 provides lightweight efficiency; ResNet50 offers higher accuracy.

3. **TensorFlow / Keras:**
    - Open-source deep learning frameworks for model design, training, and evaluation.
    - Provide GPU acceleration, extensive APIs, and visualization tools (TensorBoard).

4. **OpenCV:**
    - Used for image preprocessing and real-time video capture integration.
    - Enables image resizing, color correction, and object detection visualization.

5. **Google Colab / Jupyter Notebook:**
    - Cloud-based or local environments for training and experimentation.
    - Enable reproducibility and collaborative work among team members.

Relevance to Project

These technologies together form a robust ecosystem for developing a waste classification model:

- **CNN + Transfer Learning:** Efficiently handles complex waste images.
- **TensorFlow/Keras:** Streamlines training and hyperparameter tuning.
- **OpenCV:** Facilitates integration into a real-time camera system for live waste sorting.
- **Google Colab:** Provides accessible GPU support for model experimentation.

Technology	Strengths	Weaknesses	Suitability
CNN	Excellent for spatial features	Requires large data	High
MobileNetV2	Lightweight, fast	Slightly lower accuracy	For mobile deployment
ResNet50	Very accurate	High computational cost	For high-accuracy model
TensorFlow/Keras	Scalable, open-source	Requires GPU	For training pipeline
OpenCV	Real-time capable	Limited deep model support	For integration and preprocessing

Comparison and EvaluationUse Cases and Examples

- **Paper 2 (2025)** used a lightweight CNN and achieved real-time hardware sorting, showing the success of transfer learning models like MobileNet in waste management applications.

- **Paper 1 (2024)** utilized a CNN in TensorFlow and OpenCV for waste image preprocessing and classification, achieving over 90% accuracy.

These demonstrate the practical success of our chosen stack in similar research contexts.

Conclusion

The combination of CNN architectures, transfer learning models (MobileNetV2 and ResNet50), and frameworks like TensorFlow and OpenCV provides a powerful foundation for developing a scalable and efficient waste classification system. These technologies will enable the system to achieve high accuracy, real-time performance, and adaptability to future extensions such as object detection or robotic integration.

References

1. Khetarpal, S., & Khetarpal, A. (2024). *Trash Detection: Advanced Classification of Waste Materials Using ML Techniques*. IEEE Xplore. https://ieeexplore.ieee.org/document/10688146

2. Islam, M. R., & Roy, M. (2025). *An Automated Waste Classification System Using Deep Learning Techniques: Toward Efficient Waste Recycling and Environmental Sustainability*. Knowledge-Based Systems, Elsevier. https://www.sciencedirect.com/science/article/pii/S0950705125000760

3. Kaggle. (2024). *Waste Segregation Image Dataset*. https://www.kaggle.com/datasets/aashidutt3/waste-segregation-image-dataset/data

4. Kaggle. (2024). *Waste Classification Dataset*. https://www.kaggle.com/datasets/phenomsg/waste-classification

5. Kaggle. (2024). *Garbage Classification Dataset*. *https://www.kaggle.com/datasets/mostafaabla/garbage-classification*

6. Roboflow Universe. (2024). *Waste Object Detection Datasets*. https://universe.roboflow.com/search?q=class%3Awaste