ECS 171 Group Project 1-Pager

Group 27
Jackson Cmelak, Brian Li, Meghana Manepalli, Susanna Mathew

Note: Group leader changed to Meghana Manepalli.

I. PROBLEM STATEMENT

The average citizen of a developed country produces roughly 1,000 pounds of waste annually, highlighting the importance of the waste management industry. Since waste management facilities identify, evaluate, and manage a vast range of materials, it is imperative for them to be efficient and flexible. Unfortunately, current waste management systems are outdated [1], leading to the widespread, disorganized disposal of this waste, and posing a significant threat to the environment, public health, and resource sustainability. The far-reaching consequences of this mismanaged waste include land pollution, contamination of the water supply, and depletion of natural resources.

The central problem in waste management is properly sorting the many types of common waste. Existing methods of waste segregation that rely on human judgment are error-prone and inconsistent, leading to improper disposal and inadequate recycling rates. To address this issue, our project aims to develop a streamlined machine learning approach to automatically sort the most common types of refuse.

If robust and accurate classification is successfully implemented, it can be used in many different fields. Some examples include:

- Waste Management and Recycling
    - The model can be employed in waste sorting facilities to automate the categorization of waste into trash, compost, recyclables, and their subgroups.
- Environmental Conservation
    - By identifying and quantifying waste in the environment, the model can monitor and model pollution levels, especially air, water, and land pollution.
- Urban Planning
    - Cities can optimize waste collection routes by using the model to track the amount and type of waste in different areas, leading to more efficient garbage collection systems.

II. DATASET DESCRIPTION

The dataset selected for this project is a robust repository of thousands of images featuring different materials with various colors, patterns, lightings, and points of view.

- **Creator:** C. Chang on Kaggle
- **Title:** 'Garbage Classification'
- **Details:** 6 categories; Cardboard (403 images), Glass (501 images), Metal (410 images), Paper (594 images), Plastic (482 images), and Misc. Trash (137 images).

- **Format:** Images are kept separated in material-specific folders, along with text files containing pre-selected training and testing subsets. These subsets contain lists of the image names, along with numbers pertaining to what type of material the image contains.

III. GOALS

Our project aims to develop a machine learning-based garbage classification system with the following key objectives:

- **Data Preparation:** Collect and preprocess a diverse data set of waste material images including cardboard, glass, metal, paper, plastic, and trash. Ensure data cleanliness, balance, and integrity to facilitate model training.
- **Model Development:** Implement machine learning and deep learning algorithms, such as convolutional neural networks (CNNs), to create a robust classification model.
- **Evaluation:** Rigorously test and evaluate model performance using metrics such as accuracy, precision, recall, and F1-score. Utilize cross-validation to assess the model's generalization capability and identify potential overfitting or underfitting issues.
- **User Interface:** Develop an intuitive web-based interface that allows users to upload images of waste materials. The interface should invoke the classification model and provide real-time predictions on the type of waste, enhancing user accessibility.
- **Documentation:** Create comprehensive documentation detailing the data collection process, model architecture, hyperparameters, and performance results. This documentation will serve as a valuable resource for future machine learning practitioners and researchers.

IV. PROJECT TIMELINE

A. Description and Background: 10/16 - 10/20

- Complete "Introduction and Background" rough draft.
- Be able to describe the problem, significance, and methodology for a garbage classifier.

B. Literature Review and EDA: 10/20 - 10/27

- Complete "Literature Review" and "Dataset Description and Exploratory Data Analysis" rough draft.
- Be able to describe different models and their pros and cons.
- Work on feature engineering, selection, and extraction necessary, as well as use data visualization and statistical techniques to better understand our dataset.

C. *Building Model and Fine-tuning: 11/03 - 11/17*

- Complete "Proposed Methodology and Experimental Results" rough draft.
- Build model(s), functions for validation, and continue to finetune model for accuracy.

D. *Quality Assurance 11/17 - 11/24*

- Evaluate and test the model thoroughly.
- Begin building the webpage.

E. *Front End and Finishing Touches: 11/24 - 12/01*

- Complete the webpage.
- Complete "Conclusion and Discussion" section, ensure that "References" are complete.
- Complete any finishing touches on webpage, model, and project report.

REFERENCES

[1] Jouhara, H., Czajczyńska, D., Ghazal, H., Krzyżyńska, R., Anguilano, L., Reynolds, A. J., and Spencer, N. (2017). Municipal waste management systems for domestic use. Energy, 139, 485-506. doi: https://doi.org/10.1016/j.energy.2017.07.162