ML-Driven Waste Classification for Effective Organic and Non-Organic Waste Management

K. Hima Bindu[1], Digumarthi Nandeesh[2], Prasanjan Biswal[2], V Dheeraj Rajulu[2], Chikka Nikhil[2], Thumu Manas Anoop[2]

[1]Assistant Professor,[2]UG Student,[1,2]Department of Computer Science and Engineering (AI & ML)
[1,2]Malla Reddy Engineering College and Management Science, Kistapur, Medchal-50140l, Hyderabad, Telangana, India

Abstract

Smart Waste Collection system can be developed by optimizing waste collection routes based on real-time waste classification and reducing operational costs. Accurate waste classification enables efficient recycling practices by diverting organic waste for composting and converting non-organic waste into recyclable materials. Proper waste classification helps prevent the contamination of soil, water bodies, and air, reducing the adverse environmental impacts of mismanaged waste. By segregating organic waste for composting, valuable nutrients can be returned to the soil, promoting sustainable agriculture and conserving resources. Conventional waste classification methods often rely on manual sorting or basic rule-based systems, which are labour-intensive, time-consuming, and error-prone. Human involvement in the sorting process can lead to inconsistencies and variations in waste categorization. Rule-based systems lack the ability to handle complex and diverse waste compositions, leading to suboptimal accuracy, especially in cases of mixed waste. Moreover, these methods might not be scalable or adaptable to handle large-scale waste classification demands in urban areas. The proposed machine learning (ML)-driven waste classification system leverages the power of AI algorithms to automate and improve the waste classification process. The system employs image analysis techniques to extract visual features from waste images, such as color, texture, and shape.

Keywords: Waste management, Waste classification, Organic waste, non-organic waste, Machine learning, Predictive analytics, Random Forest Classifier.

1. INTRODUCTION

The research topic, "ML-Driven Waste Classification for Effective Organic and Non-Organic Waste Management," stands at the forefront of addressing one of the world's pressing environmental challenges: efficient waste management. As urbanization accelerates and global populations burgeon, waste generation has reached unprecedented levels, straining our ecosystems and natural resources. In this context, this research harnesses the power of Machine Learning (ML) to revolutionize waste management practices by automating the classification of waste into organic and non-organic categories [1]. The motivation behind this research is grounded in the urgent need to develop sustainable waste management solutions that mitigate environmental degradation, reduce landfill waste, and optimize resource utilization. Conventional waste sorting methods often rely on manual labour and human judgment, which are not only time-consuming but also prone to errors [2]. This research addresses these limitations by leveraging ML algorithms to analyze and classify waste items based on their composition, characteristics, and recyclability. To achieve this goal, the research delves into the development and training of ML models capable of processing images, sensor data, or other inputs to distinguish between organic waste (such as food scraps and yard trimmings) and non-organic waste (including plastics, metals, and glass). The outcome is an automated waste classification system that enhances waste sorting efficiency, enabling municipalities, recycling facilities, and individuals to manage waste streams more effectively [3].

Furthermore, the research emphasizes the ethical dimension of technology deployment. It underscores the importance of responsible AI usage, data privacy protection, and sustainability in waste management practices to ensure that the benefits of ML-driven waste classification are aligned with environmental stewardship and ethical considerations [4]. In this introductory overview, we will delve into this research's key components and objectives. We will explore the challenges posed by escalating waste generation, introduce the role of ML in waste classification, and underline the transformative potential of this research in optimizing waste management strategies. Additionally, we will highlight the ethical considerations and real-world applications of this research, which extend across municipal waste management, recycling facilities, and sustainable urban planning [5]. The "ML-Driven Waste Classification for Effective Organic and Non-Organic Waste Management" signifies a pioneering effort to harness the capabilities of ML in addressing the global challenge of waste management [6]. By automating waste classification processes, this research aims to enhance resource recovery, reduce environmental impact, and promote sustainable waste management practices while adhering to ethical standards and responsible technology use.

The research on "ML-Driven Waste Classification for Effective Organic and Non-Organic Waste Management" is motivated by a confluence of critical factors that underscore the urgent need for transformative solutions in waste management practices. First and foremost, the escalating magnitude of waste generation in our modern world serves as a compelling motivation. Rapid urbanization, population growth, and increased consumption have led to an unprecedented surge in waste production, straining existing waste management systems to their limits [7]. This surge not only poses environmental and logistical challenges but also highlights the inefficiency of traditional waste sorting methods, which are often labour-intensive, time-consuming, and prone to errors. Moreover, the pressing environmental impact of inefficient waste management practices propels this research. The environmental consequences of improper waste disposal, including overflowing landfills and uncontrolled waste incineration, are profound [8]. They contribute to the release of harmful greenhouse gases, soil and water contamination, and air pollution, thus exacerbating the global environmental crisis. The research seeks to address these challenges by harnessing Machine Learning (ML) technology to optimize waste classification, with the aim of reducing environmental degradation and promoting sustainable waste management practices [9].

Another significant motivation lies in the quest for resource optimization and recycling efficiency. Non-organic waste, which includes materials such as plastics, metals, and glass, often contains valuable resources that can be reclaimed and reused. Effective waste classification through ML-driven automation not only improves the recovery of these resources but also facilitates their integration into the circular economy, reducing the need for virgin resource extraction and conserving natural resources. This resource-centric approach aligns with sustainability goals and contributes to the responsible stewardship of our planet's resources. Furthermore, the ethical dimension of responsible waste management is a central motivation [10].

2. LITERATURE SURVEY

Fogarassy, et al. [11] proposed Composting Strategy Instead of Waste-to-Energy in the Urban Context. The objective of this work is to identify the barriers to organic waste management solutions from an actor's perspective and to explore their causal relationships to overcome the organic waste management problem from a system perspective. Several key challenges were identified regarding organic waste management solutions, the current intervention overview indicates that promoting and tracking attention towards “value to waste” would be an effective solution approach. Kharola, et al. [12] proposed Barriers to organic waste management in a circular economy. The objective of this study is to identify the barriers to organic waste management solutions from an actor's perspective and to explore their causal relationships to overcome the organic waste management problem from a system perspective. Several key challenges were identified regarding organic waste management solutions, the current intervention overview indicates that promoting and tracking attention towards "value to waste" would be an effective solution approach.

Loganayagi, et al. [13] proposed An Automated Approach to Waste Classification Using Deep Learning. The study developed a custom inception model by adding additional layers and compares the performance through accuracy against the basic Inceptionv3 model. The study used SGD (stochastic gradient descent) with liner regression algorithm for classification and categorical cross-entropy for loss estimation. The current study uses the ReLU function to overcome the under-fitting and over-fitting issues. Mookkaiah, et al. [14] proposed the Design and development of a smart Internet of Things–based solid waste management system using computer vision. The proposed model identifies the type of waste and classifies them as biodegradable or non-biodegradable to collect in respective waste bins precisely. Furthermore, observation of performance metrics, accuracy, and loss ensures the effective functions of the proposed model compared to other existing models. The proposed ResNet-based CNN performs waste classification with 19.08% higher accuracy and 34.97% lower loss than the performance metrics of other existing models.

Alvianingsih, et al. [15] proposed an Automatic garbage classification system using arduino-based controller and binary tree concept. The proposed design consists of an automatic door, garbage sorter, user interface, and capacity observer. The main components of the system are Microcontroller Arduino Mega 2560, ultrasonic sensor HCSR04, servo motor MG996R, Inductive Proximity Sensor, and Capacitive Proximity Sensor. From the performance test result we can obtain that HC-SR04 ultrasonic sensor as an object detector has an error in distance stabilization of 33.3%, inductive proximity sensors as metal detectors have a 100 % success rate, while capacitive proximity sensors as organic garbage detector has a success rate of 85.7 %.

3. PROPOSED SYSTEM

3.1 Overview

This project focuses on image classification, specifically distinguishing between organic and non-organic objects. It begins with image preprocessing to prepare the data, followed by dataset splitting to create training and testing subsets. The Random Forest Classifier is chosen as the classification model, and it undergoes thorough examination for accuracy, precision, recall, and overall readiness for deployment. This project's goal is to create a robust image classification model capable of accurately identifying organic and non-organic objects, with potential applications in fields such as agriculture, waste management, and environmental monitoring. Figure 4.1 shows the proposed system model. The detailed operation illustrated as follows:

step 1: Image Preprocessing: The project begins by collecting an extensive range of images that contain both organic and non-organic objects. Image preprocessing is the initial step, where the collected images undergo various transformations and enhancements to prepare them for analysis. Preprocessing steps may include resizing, cropping, color normalization, and noise reduction.

step 2: Dataset Splitting: After preprocessing the images, the dataset is divided into two subsets: a training set and a testing set. The common split ratio is 80% for training and 20% for testing. This division ensures that the model is trained on a substantial portion of the data while also reserving a separate portion for evaluating its performance.

step 3: RFC Training Model:

- A Random Forest Classifier (RFC) model is chosen for the image classification task. RFC is an ensemble learning method known for its versatility and effectiveness in handling image data.
- The training phase involves feeding the preprocessed images from the training set into the RFC model. During this process, the model learns patterns and features that distinguish organic and non-organic objects. It aims to create a decision boundary that can accurately classify images based on their content.

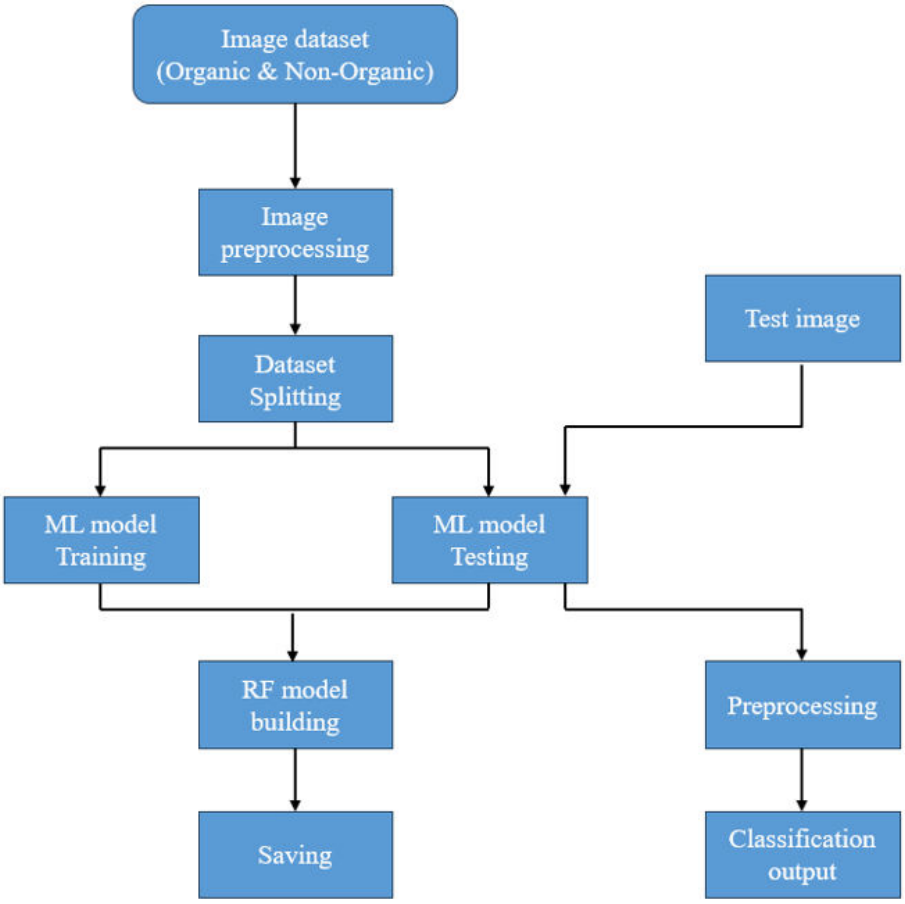

Figure 1: Proposed System model.

step 4: Model Examination: After training, the RFC model is subjected to a comprehensive examination to assess its performance in classifying organic and non-organic images. This evaluation typically includes several key aspects:

- **Accuracy**: This metric measures the percentage of correctly classified images out of the total number of images in the testing set. It provides an overall assessment of how well the model is performing.
- **Precision**: Precision measures the proportion of true positive predictions (correctly classified organic images) out of all the positive predictions (organic images). It quantifies the model's ability to avoid false positives.
- **Recall (Sensitivity)**: Recall calculates the proportion of true positive predictions out of all actual organic images in the testing set. It reflects the model's ability to detect organic objects accurately.

- **F1-Score**: The F1-Score is the harmonic mean of precision and recall. It offers a balanced assessment of the model's performance, especially when class distribution is imbalanced.

3.2 Data Preprocessing

Image preprocessing is a critical step in computer vision and image analysis tasks. It involves a series of operations to prepare raw images for further processing by algorithms or neural networks. Here's an explanation of each step in image preprocessing:

Step 0. Image Read: The first step in image preprocessing is reading the raw image from a source, typically a file on disk. Images can be in various formats, such as JPEG, PNG, BMP, or others. Image reading is performed using libraries or functions specific to the chosen programming environment or framework. The result of this step is a digital representation of the image that can be manipulated programmatically.

1. Image Resize: Image resizes is a common preprocessing step, especially when working with machine learning models or deep neural networks. It involves changing the dimensions (width and height) of the image. Resizing can be necessary for several reasons:

- Ensuring uniform input size: Many machine learning models, especially convolutional neural networks (CNNs), require input images to have the same dimensions. Resizing allows you to standardize input sizes.
- Reducing computational complexity: Smaller images require fewer computations, which can be beneficial for faster training and inference.
- Managing memory constraints: In some cases, images need to be resized to fit within available memory constraints.

When resizing, it's essential to maintain the aspect ratio to prevent image distortion. Typically, libraries like OpenCV or Pillow provide convenient functions for resizing images.

2. Image to Array: In this step, the image is converted into a numerical representation in the form of a multidimensional array or tensor. Each pixel in the image corresponds to a value in the array. The array is usually structured with dimensions representing height, width, and color channels (if applicable).

For grayscale images, the array is 2D, with each element representing the intensity of a pixel. For color images, it's a 3D or 4D array, with dimensions for height, width, color channels (e.g., Red, Green, Blue), and potentially batch size (if processing multiple images simultaneously).

The conversion from an image to an array allows for numerical manipulation and analysis, making it compatible with various data processing libraries and deep learning frameworks like NumPy or TensorFlow.

3. Image to Float32: Most machine learning and computer vision algorithms expect input data to be in a specific data type, often 32-bit floating-point numbers (float32). Converting the image array to float32 ensures that the pixel values can represent a wide range of intensities between 0.0 (black) and 1.0 (white) or sometimes between -1.0 and 1.0, depending on the specific normalization used.

This step is essential for maintaining consistency in data types and enabling compatibility with various machine learning frameworks and libraries. It's typically performed by dividing the pixel values by the maximum intensity value (e.g., 255 for an 8-bit image) to scale them to the [0.0, 1.0] range.

4. Image to Binary: Image binarization is a process of converting a grayscale image into a binary image, where each pixel is represented by either 0 (black) or 1 (white) based on a specified threshold. Binarization is commonly used for tasks like image segmentation, where you want to separate objects from the background.

The process involves setting a threshold value, and then for each pixel in the grayscale image, if the pixel value is greater than or equal to the threshold, it is set to 1; otherwise, it is set to 0.

Binarization simplifies the image and reduces it to essential information, which can be particularly useful in applications like character recognition or object tracking, where you need to isolate regions of interest.

3.3 Dataset Splitting

In machine learning data pre-processing, we divide our dataset into a training set and test set. This is one of the crucial steps of data pre-processing as by doing this, we can enhance the performance of our machine learning model. Suppose if we have given training to our machine learning model by a dataset and we test it by a completely different dataset. Then, it will create difficulties for our model to understand the correlations between the models. If we train our model very well and its training accuracy is also very high, but we provide a new dataset to it, then it will decrease the performance. So we always try to make a machine learning model which performs well with the training set and also with the test dataset. Here, we can define these datasets as:

Training Set: A subset of dataset to train the machine learning model, and we already know the output.

Test set: A subset of dataset to test the machine learning model, and by using the test set, model predicts the output.

4. RESULTS AND DISCUSSION

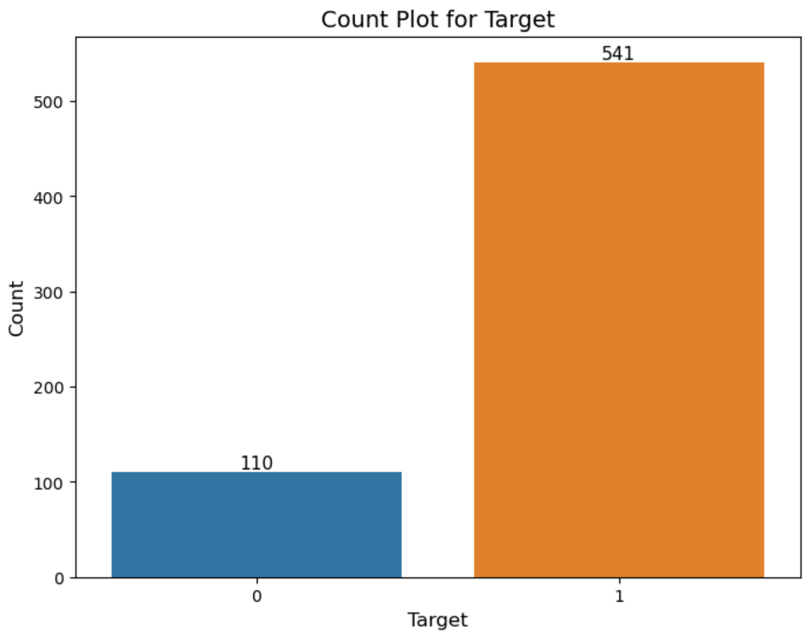

Figure 2: Count plot for target column in a dataset

...	67490	67491	67492	67493	67494	67495	67496	67497	67498	67499
...	0.399986	0.498039	0.462745	0.396078	0.494131	0.458837	0.392170	0.494118	0.458824	0.392157
...	0.608712	0.742601	0.680832	0.577866	0.758485	0.701617	0.591800	0.732052	0.666371	0.566103
...	0.453915	0.468621	0.452934	0.456856	0.495102	0.481378	0.485300	0.580346	0.572502	0.576424
...	1.000000	0.999018	1.000000	0.998527	0.999017	1.000000	0.998035	0.999754	1.000000	0.999508
...	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000
...	...	...	...	...	...	...	...	...	...	...
...	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000
...	0.356485	0.620515	0.577060	0.492773	0.644246	0.581459	0.510193	0.612403	0.518076	0.430977
...	0.406301	0.388129	0.428749	0.387117	0.399592	0.409507	0.388051	0.400000	0.407843	0.388235
...	0.282353	0.196078	0.200000	0.282353	0.196078	0.200000	0.282353	0.196078	0.200000	0.282353
...	0.007844	0.007843	0.007843	0.007843	0.007843	0.007843	0.007843	0.007843	0.007843	0.007843

Figure 3: Data frame of image data after preprocessing

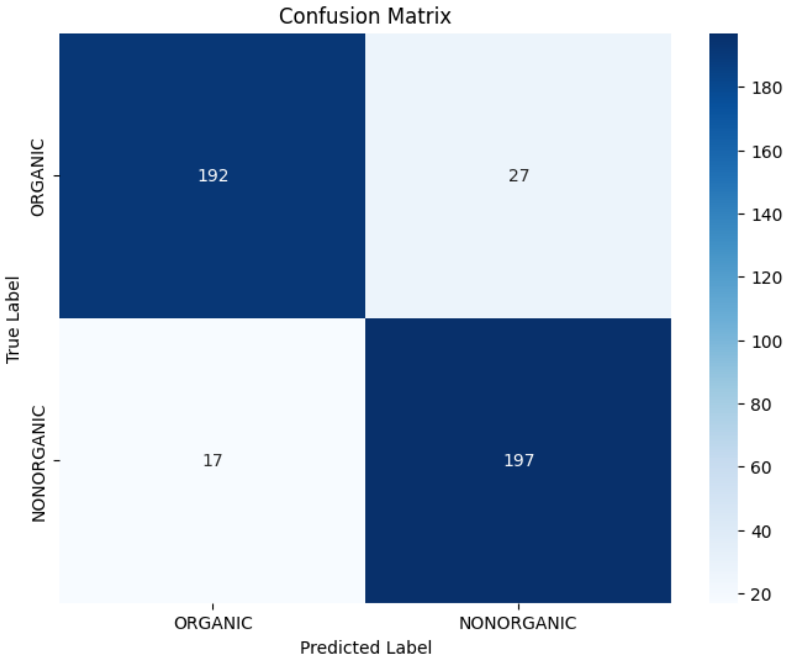

Figure 4: Heatmap of confusion matrix for Random Forest algorithm

```
classification_report:
               precision    recall  f1-score   support

           0       0.92      0.88      0.90       219
           1       0.88      0.92      0.90       214

    accuracy                           0.90       433
   macro avg       0.90      0.90      0.90       433
weighted avg       0.90      0.90      0.90       433
```

Figure 5: Classification report of Random Forest classifier

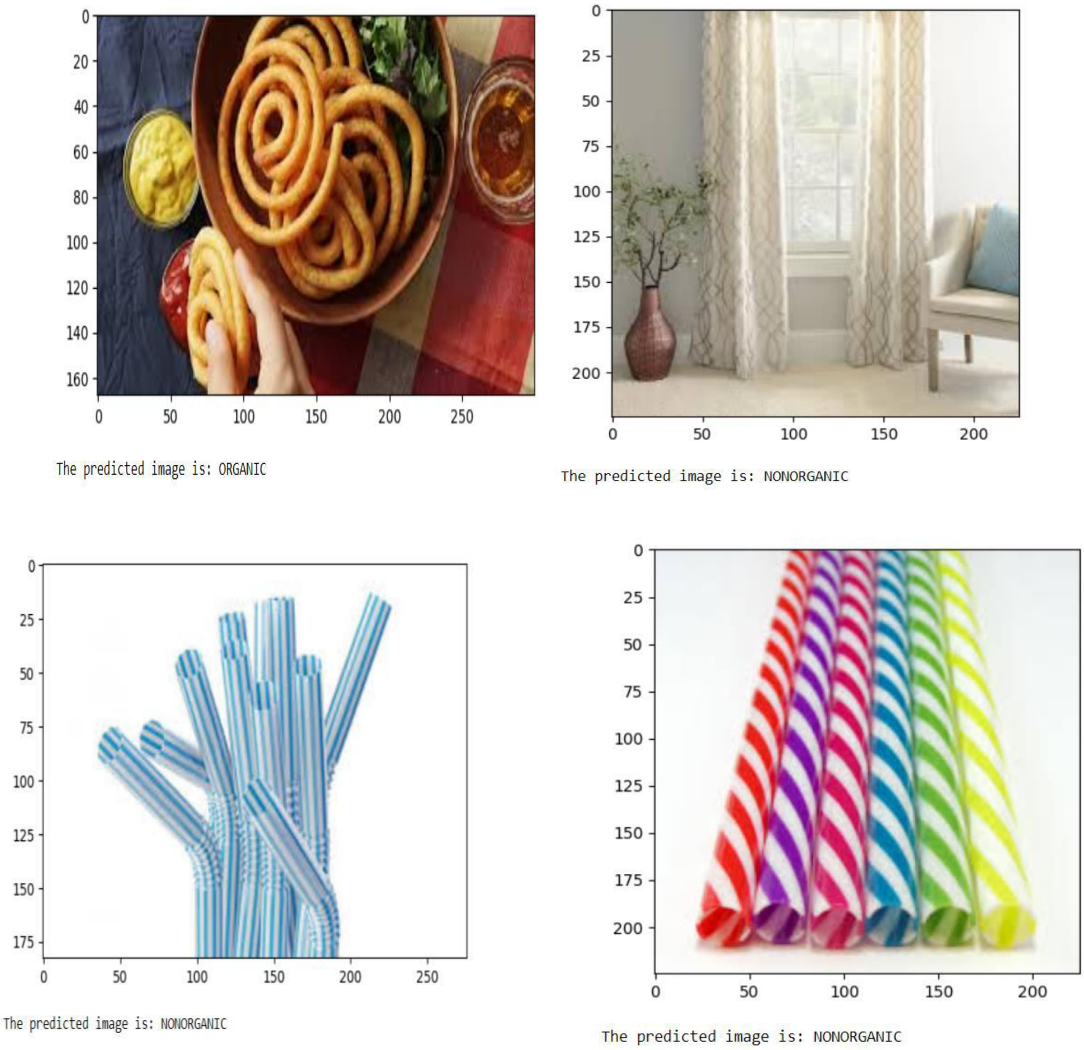

Figure 6: Prediction results on test data using random forest classifier

5. Conclusion

The project, "Image Classification for Organic and Non-Organic Objects," has successfully demonstrated a systematic workflow for differentiating between organic and non-organic images. Beginning with extensive data collection and image preprocessing, the project prepared the dataset for model training and evaluation. The dataset was thoughtfully divided into a training set and a testing set, with an 80-20 split ratio. A Random Forest Classifier (RFC) was employed to train on the preprocessed images, learning to distinguish between organic and non-organic objects based on image features. Model examination yielded insights into its performance, with accuracy, precision, recall, and the F1-score providing comprehensive metrics for evaluation. This project signifies a crucial step in automating the classification of images, with potential applications in industries such as agriculture, waste management, and environmental monitoring, where distinguishing between organic and non-organic materials is essential for decision-making and resource management.

REFERENCES

[1] Lingaraju, Abhishek Kadalagere, et al. "IoT-Based Waste Segregation with Location Tracking and Air Quality Monitoring for Smart Cities." Smart Cities 6.3 (2023): 1507-1522.

[2] Swarnawati, Aminah, K. N. Jamiati, and Robby Milana. "ROLE OF WASTE BANK IN DISSEMINATION ENVIRONMENT CARE MESSAGE." Proceedings Of International Conference On Communication Science. Vol. 2. No. 1. 2022.

[3] Aziz, Abdul, M. Imam Tobroni, and Suprayitno Sutoyo. "Environmental Awareness Education Campaign through Videography at the Rawajati Waste Bank in South Jakarta." KnE Social Sciences (2023): 31-39.

[4] Laksmono, Bambang Shergi. "Implementation of Creating Shared Value at PT Pegadaian (Persero) Case Study of Alamanda Sejahtera Waste Bank Assistance Program in Bekasi City." Devotion Journal of Community Service 4.6 (2023): 1289-1296.

[5] Shah, Rajiv V., and Samapti Guha. "Waste management and private sector participation-operational and behavioural perspectives." International Journal of Environment and Waste Management 32.1 (2023): 113-127.

[6] Baojun, Gong, et al. "AI-based detection system of resident's behaviors in automatic trash sorting booths: a background computing-based solution." 2022 China Automation Congress (CAC). IEEE, 2022.

[7] Angraini, Peggy Dian Septi Nur. "Legal effectiveness of independent waste management through waste banks in the City of Tegal as an effort to empower and care for the environment." The International Journal of Politics and Sociology Research 11.2 (2023): 195-204.

[8] Adriansyah, Endi, et al. "Decreasing pH, COD and TSS of Domestic Liquid Waste Using Photocatalysis TiO2 (Titanium Dioxide)." International Journal of Research in Vocational Studies (IJRVOCAS) 3.2 (2023): 11-15.

[9] Nguyen, Trang Thi Thu, et al. "Household food waste disposal behaviour is driven by perceived personal benefits, recycling habits and ability to compost." Journal of Cleaner Production 379 (2022): 134636.

[10] Kumar, Akhilesh, and Avlokita Agrawal. "Pneumatic waste collection system-A proposed innovation for Indian Institute of Technology, Roorkee Campus, India." IOP Conference Series: Earth and Environmental Science. Vol. 1210. No. 1. IOP Publishing, 2023.

[11] Fogarassy, Csaba, Nguyen Huu Hoang, and Kinga Nagy-Pércsi. "Composting Strategy Instead of Waste-to-Energy in the Urban Context—A Case Study from Ho Chi Minh City, Vietnam." Applied Sciences 12.4 (2022): 2218.

[12] Kharola, Shristi, et al. "Barriers to organic waste management in a circular economy." Journal of Cleaner Production 362 (2022): 132282.

[13] Loganayagi, S., and D. Usha. "An Automated Approach to Waste Classification Using Deep Learning." 2023 Fifth International Conference on Electrical, Computer and Communication Technologies (ICECCT). IEEE, 2023.

[14] Mookkaiah, Senthil Sivakumar, et al. "Design and development of smart Internet of Things–based solid waste management system using computer vision." Environmental Science and Pollution Research 29.43 (2022): 64871-64885.

[15] Alvianingsih, Ginas, Tri Wahyu Oktaviana Putri, and Danu Azhar Hidayat. "Automatic garbage classification system using arduino-based controller and binary tree concept." 2022 11th Electrical Power, Electronics, Communications, Controls and Informatics Seminar (EECCIS). IEEE, 2022.